

PT preventing pneumococcal infections, particularly infections caused by
 PT Streptococcus, e.g. otitis media, nasopharyngeal, bronchial, lung or
 PT blood infections -

XX Claim 4; Page 47-50; 54pp; English.

XX The invention relates to novel immunogenic polypeptides, Spi28 and Spi30
 CC from *S. pneumoniae*. Vaccines comprising the polypeptides are useful for
 CC the treatment and prevention of pneumococcal infections, particularly
 CC infections caused by *Streptococcus*, such as otitis media, nasopharyngeal,
 CC bronchial, lung or blood infections. The antigens are used as immunogenic
 CC agents to stimulate an immune response. The antisera and antibodies may
 CC also be used in diagnosing and treating pneumococcal infections.
 CC Recombinant polypeptides serve as a mechanism for stimulating production
 CC of antibodies for use in passive immunotherapy, diagnostic reagents, and
 CC as reagents in other processes such as affinity chromatography. The
 CC present sequence represents the *S. pneumoniae* Spi28 polypeptide.

XX Sequence 664 AA;

Query Match 100.0%; Score 4398; DB 22; Length 664;

Best Local Similarity 100.0%; Pred. No. 2.3e-77; Mismatches 0; Indels 0; Gaps 0;

Matches 664; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YPVVLADTSSSEDALNIDKEKVAENKEKHENIHSAHETSODPFKEKTAIVIKEKEVSKN 60
 DB 1 YPVVLADTSSSEDALNIDKEKVAENKEKHENIHSAHETSODPFKEKTAIVIKEKEVSKN 60

QY 61 PVIDNNTSNEAKIKEENSKNSQGDYTDSPVNKNTENPKKEDKVYIAEFKDGESGEKAI 120
 DB 61 PVIDNNTSNEAKIKEENSKNSQGDYTDSPVNKNTENPKKEDKVYIAEFKDGESGEKAI 120

QY 61 PVIDNNTSNEAKIKEENSKNSQGDYTDSPVNKNTENPKKEDKVYIAEFKDGESGEKAI 120
 DB 61 PVIDNNTSNEAKIKEENSKNSQGDYTDSPVNKNTENPKKEDKVYIAEFKDGESGEKAI 120

QY 121 KELSLKNTKVLTYTDRIFNHSAIETTPDNLDKIKQIEGISVERAKOVPMNNHARKEI 180
 DB 121 KELSLKNTKVLTYTDRIFNHSAIETTPDNLDKIKQIEGISVERAKOVPMNNHARKEI 180

QY 121 KELSLKNTKVLTYTDRIFNHSAIETTPDNLDKIKQIEGISVERAKOVPMNNHARKEI 180
 DB 121 KELSLKNTKVLTYTDRIFNHSAIETTPDNLDKIKQIEGISVERAKOVPMNNHARKEI 180

QY 181 GVEBAIDVLSKINAFPGKDFGRGVNISNIDTGTDRYRHKAMRIDDAVASRPFKEDLKG 240
 DB 181 GVEBAIDVLSKINAFPGKDFGRGVNISNIDTGTDRYRHKAMRIDDAVASRPFKEDLKG 240

QY 181 GVEBAIDVLSKINAFPGKDFGRGVNISNIDTGTDRYRHKAMRIDDAVASRPFKEDLKG 240
 DB 181 GVEBAIDVLSKINAFPGKDFGRGVNISNIDTGTDRYRHKAMRIDDAVASRPFKEDLKG 240

QY 241 TDKNYMSLDKIPHANVYNGGKITTEKTDGDRYDFPHGMHAGLADNDTEODIKNENG 300
 DB 241 TDKNYMSLDKIPHANVYNGGKITTEKTDGDRYDFPHGMHAGLADNDTEODIKNENG 300

QY 241 TDKNYMSLDKIPHANVYNGGKITTEKTDGDRYDFPHGMHAGLADNDTEODIKNENG 300
 DB 241 TDKNYMSLDKIPHANVYNGGKITTEKTDGDRYDFPHGMHAGLADNDTEODIKNENG 300

QY 301 IDGAPNAQIFSYKMYSDAGSGFADDETFPHAIEDSIRKNDVSVSSGFGTGLVGEKY 360
 DB 301 IDGAPNAQIFSYKMYSDAGSGFADDETFPHAIEDSIRKNDVSVSSGFGTGLVGEKY 360

QY 301 IDGAPNAQIFSYKMYSDAGSGFADDETFPHAIEDSIRKNDVSVSSGFGTGLVGEKY 360
 DB 301 IDGAPNAQIFSYKMYSDAGSGFADDETFPHAIEDSIRKNDVSVSSGFGTGLVGEKY 360

QY 361 WQAIRALKRAGIPNVVATGNVATSSASSSSWDLVANNHLMKTDGNTVTRTAHEDAIIVAS 420
 DB 361 WQAIRALKRAGIPNVVATGNVATSSASSSSWDLVANNHLMKTDGNTVTRTAHEDAIIVAS 420

QY 361 WQAIRALKRAGIPNVVATGNVATSSASSSSWDLVANNHLMKTDGNTVTRTAHEDAIIVAS 420
 DB 361 WQAIRALKRAGIPNVVATGNVATSSASSSSWDLVANNHLMKTDGNTVTRTAHEDAIIVAS 420

QY 421 AKNQVBERDKVNIIGESFRYRNIGAFPKSKITTNEDGTAPSKLKFYIIGKQDQDLIG 480
 DB 421 AKNQVBERDKVNIIGESFRYRNIGAFPKSKITTNEDGTAPSKLKFYIIGKQDQDLIG 480

QY 421 AKNQVBERDKVNIIGESFRYRNIGAFPKSKITTNEDGTAPSKLKFYIIGKQDQDLIG 480
 DB 421 AKNQVBERDKVNIIGESFRYRNIGAFPKSKITTNEDGTAPSKLKFYIIGKQDQDLIG 480

QY 481 LDLRGKIIVMDRIYTKDKLNAFKKAMDKGARAIWVNVVYNNRDNMTLPMAGYEADG 540
 DB 481 LDLRGKIIVMDRIYTKDKLNAFKKAMDKGARAIWVNVVYNNRDNMTLPMAGYEADG 540

QY 481 LDLRGKIIVMDRIYTKDKLNAFKKAMDKGARAIWVNVVYNNRDNMTLPMAGYEADG 540
 DB 481 LDLRGKIIVMDRIYTKDKLNAFKKAMDKGARAIWVNVVYNNRDNMTLPMAGYEADG 540

QY 541 TKSQVFSISGDDGVFLAMNINPDKTEVYRNKKEDEPKDLBOYPIIDMESFNSNKNPVND 600
 DB 541 TKSQVFSISGDDGVFLAMNINPDKTEVYRNKKEDEPKDLBOYPIIDMESFNSNKNPVND 600

QY 541 TKSQVFSISGDDGVFLAMNINPDKTEVYRNKKEDEPKDLBOYPIIDMESFNSNKNPVND 600
 DB 541 TKSQVFSISGDDGVFLAMNINPDKTEVYRNKKEDEPKDLBOYPIIDMESFNSNKNPVND 600

QY 601 EKEIDFKAPDTEKLYEDTIVPAGSTSWGPRIDLKLPDVSAGKNIKSTLVYINGKS 660
 DB 601 EKEIDFKAPDTEKLYEDTIVPAGSTSWGPRIDLKLPDVSAGKNIKSTLVYINGKS 660

QY 661 TYGY 664
 DB 661 TYGY 664

RESULT 2
 AA81710
 ID AA81710 standard; Protein; 2120 AA.

XX AA81710;

DT 02-JUN-2000 (first entry)

DE Streptococcus pneumoniae protein sequence ID3.

XX Streptococcus pneumoniae infection; immunogen; antigen; diagnosis; AIDS;
 KW bacterial pneumonia; asplenia; heart disease; lung disease; alcoholism;
 KW kidney disease; diabetes; immunosuppressive disorder; otitis media;
 KW pneumococcal septicemia; sinusitis; meningitis; therapy.

OS Streptococcus pneumoniae.

XX WO200006738-A2.

XX 10-FEB-2000.

XX 27-JUL-1999; 99WO-GB02452.

XX 27-JUL-1998; 98GB-0016336.

XX 19-MAR-1999; 99US-0125329.

PA (MICR-) MICROBIAL TECHNICS LTD.

PI Le Page RWF, Wells JM, Hamifly SB, Hansbro PM;

DR WPI; 2000-195301/17.

XX N-PSDB; AA291806.

PT Streptococcal proteins and polynucleotides useful for diagnosis,
 PT treatment and prophylaxis of bacterial infections -

XX Claim 2; Page 41-42; 76pp; English.

XX This sequence represents a *Streptococcus pneumoniae* protein of the
 CC invention. The proteins (or their homologues, derivatives and/or
 CC fragments) are useful as immunogens or antigens. Immunogenic or antigenic
 CC compositions comprising the proteins are useful for vaccines and also in
 CC diagnostic assays. The sequences are useful for the detection or
 CC diagnosis of *S. pneumoniae* infection, by contacting a sample to be tested
 CC with them. Agents capable of antagonising, inhibiting or interfering with
 CC the function or expression of the protein or polypeptide are useful in
 CC medical compositions in the treatment or prophylaxis of *S. pneumoniae*
 CC infection. As the sequences can be used to treat *S. pneumoniae* infection,
 CC they can be used to treat bacterial pneumonia, which has high rates in
 CC young children, the elderly, and in patients with predisposing conditions
 CC such as asplenia, heart, lung and kidney disease, diabetes, alcoholism,
 CC or with immunosuppressive disorders, especially AIDS. They can also be
 CC used to treat pneumococcal septicemia, otitis media, sinusitis, and
 CC meningitis.

XX Sequence 2120 AA;

Query Match 100.0%; Score 4398; DB 21; Length 2120;

Best Local Similarity 100.0%; Pred. No. 4.5e-75; Mismatches 0; Indels 0; Gaps 0;

Matches 664; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YPVVLADTSSSEDALNIDKEKVAENKEKHENIHSAHETSODPFKEKTAIVIKEKEVSKN 60
 DB 1 YPVVLADTSSSEDALNIDKEKVAENKEKHENIHSAHETSODPFKEKTAIVIKEKEVSKN 60

QY 61 PVIDNNTSNEAKIKEENSKNSQGDYTDSPVNKNTENPKKEDKVYIAEFKDGESGEKAI 120
 DB 61 PVIDNNTSNEAKIKEENSKNSQGDYTDSPVNKNTENPKKEDKVYIAEFKDGESGEKAI 120

QY 61 PVIDNNTSNEAKIKEENSKNSQGDYTDSPVNKNTENPKKEDKVYIAEFKDGESGEKAI 120
 DB 61 PVIDNNTSNEAKIKEENSKNSQGDYTDSPVNKNTENPKKEDKVYIAEFKDGESGEKAI 120

QY 121 KELSLKNTKVLTYTDRIFNHSAIETTPDNLDKIKQIEGISVERAKOVPMNNHARKEI 180
 DB 121 KELSLKNTKVLTYTDRIFNHSAIETTPDNLDKIKQIEGISVERAKOVPMNNHARKEI 180

```

QY 181 GVEBAIDYLSINAPFGKNPDRGMVTSNIDTGTDRHKAMRIDDDAKASMRFKEDLKG 240
Db 181 GVEBAIDYLSINAPFGKNPDRGMVTSNIDTGTDRHKAMRIDDDAKASMRFKEDLKG 240
QY 241 TDKRYWLSDKIPHAFFNYNGSKITVEKYDDGRDYPDPHGMHIAIGLGNDEBODIKNPN 300
Db 241 TDKRYWLSDKIPHAFFNYNGSKITVEKYDDGRDYPDPHGMHIAIGLGNDEBODIKNPN 300
QY 301 IDGIAPNAQIFSYKMYSDAGSGFAGDETMFAIEDSIKHNVDVYVSSGFTGTGLVGEKY 360
Db 301 IDGIAPNAQIFSYKMYSDAGSGFAGDETMFAIEDSIKHNVDVYVSSGFTGTGLVGEKY 360
QY 361 WQAIRALRKAGIPMVVATGNVATSSASSSWDLVANNHLMKTDGNTVTRTAHEDAIAVAS 420
Db 361 WQAIRALRKAGIPMVVATGNVATSSASSSWDLVANNHLMKTDGNTVTRTAHEDAIAVAS 420
QY 421 AKONTVEFDKNIGSGESFKYRNIGAFPDKSKITTNEGDTKPSLTKKVVYIGKGDODLIG 480
Db 421 AKONTVEFDKNIGSGESFKYRNIGAFPDKSKITTNEGDTKPSLTKKVVYIGKGDODLIG 480
QY 481 LDLRGKIAVMDRIYTKDLKNAFKKAMDGAIAIMVNTVYNNRDNNTLPRMGYEADDEG 540
Db 481 LDLRGKIAVMDRIYTKDLKNAFKKAMDGAIAIMVNTVYNNRDNNTLPRMGYEADDEG 540
QY 541 TKSQVFSISGDDGYKLMNMIPDKTEYKRNKEDFPDKLEQYYPIMESFNSKPNVGD 600
Db 541 TKSQVFSISGDDGYKLMNMIPDKTEYKRNKEDFPDKLEQYYPIMESFNSKPNVGD 600
QY 601 EKEIDFPAPDTEKLYKEDIIVAGSTSWGPRIDLKLPVSAFGNISTLNVINGKS 660
Db 601 EKEIDFPAPDTEKLYKEDIIVAGSTSWGPRIDLKLPVSAFGNISTLNVINGKS 660
QY 661 TYGY 664
Db 661 TYGY 664

```

RESULT 3
ABU01020
ID ABU01020 standard; Protein; 2140 AA.

```

XX AC ABU01020;
XX DT 11-FEB-2003 (first entry)
XX DE S. pneumoniae type 4 strain protein from coding region #590.
XX KM Bacterial meningitis; pneumonia; sepsis; otitis media;
XX KM ear infection; antiinflammatory; antibacterial; immunostimulant;
XX KM auditory; respiratory; gene therapy; vaccine.
XX OS Streptococcus pneumoniae type 4 strain.
XX PN W0200277021-A2.
XX PD 03-OCT-2002.
XX PF 27-MAR-2002; 2002WO-IB02163.
XX PR 27-MAR-2001; 2001GB-0007658.
XX PA (CHIR-) CHIRON SPA.
XX PA (GENO-) INST GENOMIC RES.
XX PI Maignant V, Tettelin H, Fraser C;
XX DR WPI; 2003-040579/03.
XX DR N-PSDB; ABX06302.
XX PT New proteins and nucleic acid molecules from Streptococcus pneumoniae,
XX PT useful as medicaments for treating or preventing a disease or infection
XX PT due to streptococcus bacteria, such as pneumonia, sepsis, otitis media
XX PT or ear infection

```

```

XX XX Claim 1; SEQ ID No 1180; 56pp; English.
PS PS
XX The invention relates to a protein comprising or having at least 50%
XX identity to any of the 2469 amino acid sequences, identified in the
XX specification (available on a computer readable format), or its fragment,
XX expressed from 2469 of 2489 identified DNA coding regions from the
XX Streptococcus pneumoniae type 4 strain genomic sequence appearing as
XX AB556454. Also included are an antibody which binds one of the
XX proteins, treating a patient by administering the protein, DNA or
XX antibody (in a composition), a kit comprising first and second primers,
XX which are the nucleic acid cited above or fragments between nucleotides
XX 8-100 of a sequence not defined in the specification, for amplifying a
XX target sequence contained within a Streptococcus nucleic acid sequence,
XX where the first primer is substantially complementary to the target
XX sequence and the second primer is substantially complementary to the
XX complement of the target sequence, and where the pairs of the primers
XX having substantial complementarity define the termini of the target
XX sequence to be amplified, assay comprising contacting a test compound
XX with the protein, and determining whether the test compound binds to the
XX protein and a Streptococcus pneumoniae bacterium, where one or more
XX genes encoding the proteins has been rendered inactive. The proteins,
XX nucleic acid molecules, antibody and compositions are useful as
XX medicaments for treating or preventing a disease or infection due to
XX streptococcus bacteria, particularly S. pneumoniae, such as pneumonia,
XX sepsis, otitis media or ear infection. They are also useful in developing
XX vaccines, diagnostics and antibiotics. The methods are useful for
XX identifying immunodominant proteins. The present sequence is one of
XX the 2469 proteins expressed by the identified coding regions from the
XX genomic sequence.
XX Note: The sequence data for this patent did not form part
XX of the printed specification, but was obtained in electronic
XX format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 2140 AA;

```

Query Match 99.9%; Score 4393; DB 24; Length 2140;
Best Local Similarity 99.6%; Pred. No. 66-75;
Matches 664; Conservative 0; Mismatches 0; Indels 3; Gaps 2;

```

QY 1 Y-P--VVLADTSSSEDALNISDEKVAENKKEHNIHSAETSQDFEKKTAIVKEKEV 57
Db 19 YAPNEVVLADTSSSEDALNISDEKVAENKKEHNIHSAETSQDFEKKTAIVKEKEV 78
QY 58 SKRPVVDNNTSNEAKIKENSNSKQDYYDSFVNKNTENPKEDKRYVYIAEPDKESGE 117
Db 79 SKRPVVDNNTSNEAKIKENSNSKQDYYDSFVNKNTENPKEDKRYVYIAEPDKESGE 138
QY 118 KAIKELSSLKNTKVLTYDRIFNGSAIETTPDMLDKIQTJEGISSVERAQKQVPMNHAR 177
Db 139 KAIKELSSLKNTKVLTYDRIFNGSAIETTPDMLDKIQTJEGISSVERAQKQVPMNHAR 198
QY 178 KEIGVEBAIDYLSINAPFGKNPDRGMVTSNIDTGTDRHKAMRIDDDAKASMRFKED 237
Db 199 KEIGVEBAIDYLSINAPFGKNPDRGMVTSNIDTGTDRHKAMRIDDDAKASMRFKED 258
QY 238 LKGTDRKRYWLSDKIPHAFFNYNGSKITVEKYDDGRDYPDPHGMHIAIGLGNDEBODIKN 297
Db 259 LKGTDRKRYWLSDKIPHAFFNYNGSKITVEKYDDGRDYPDPHGMHIAIGLGNDEBODIKN 318
QY 298 FNGIDGIAPNAQIFSYKMYSDAGSGFAGDETMFAIEDSIKHNVDVYVSSGFTGTGLV 357
Db 319 FNGIDGIAPNAQIFSYKMYSDAGSGFAGDETMFAIEDSIKHNVDVYVSSGFTGTGLV 378
QY 358 EKYWQAIRALRKAGIPMVVATGNVATSSASSSWDLVANNHLMKTDGNTVTRTAHEDAIA 417
Db 379 EKYWQAIRALRKAGIPMVVATGNVATSSASSSWDLVANNHLMKTDGNTVTRTAHEDAIA 438
QY 418 VASAKQNTVEFDKNIGSGESFKYRNIGAFPDKSKITTNEGDTKPSLTKKVVYIGKGDOD 477
Db 439 VASAKQNTVEFDKNIGSGESFKYRNIGAFPDKSKITTNEGDTKPSLTKKVVYIGKGDOD 498

```

```

Qy 478 LIGDLRGKIAVMDRIYTKDLKNAFKAMDKGARAIMVNVTVNYYNRDNTLPAWGVEA 537
    |||||
Db 499 LIGDLRGRKIAVMDRIYTKDLKNAFKAMDKGARAIMVNVTVNYYNRDNTLPAWGVEA 558
Qy 538 DEGTSGQVFSISGDDGVKLMNMINPDKTEVRANKNEDFKYLEQYYPIDMESFNSNKEN 597
    |||||
Db 559 DEGTSGQVFSISGDDGVKLMNMINPDKTEVRANKNEDFKYLEQYYPIDMESFNSNKEN 618
Qy 598 VGDEKEIDPFKPAIDPKDELKYEKEDIIIVPAGSTSGPRIDLLKPDVASPGKNIKSTLNVIN 657
    |||||
Db 619 VGDEKEIDPFKPAIDPKDELKYEKEDIIIVPAGSTSGPRIDLLKPDVASPGKNIKSTLNVIN 678
Qy 658 GKSTYGY 664
    |||||
Db 679 GKSTYGY 685

```

RESULT 4
AAM61246
ID AAM61246 standard; Protein; 1007 AA.

AC AAM61246;
DT 02-OCT-1998 (first entry)

DE Streptococcus pneumoniae SPI22 protein.
KM Streptococcus pneumoniae; antigen; vaccine; infection; diagnosis;
KW detection; pneumonia; otitis media; meningitis.

XX Streptococcus pneumoniae.

XX MO9818930-A2.

XX 07-MAY-1998.

XX 30-OCT-1997; 97WO-US19422.

XX 31-OCT-1996; 96US-0029960.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Chai GH, Hromockyj A, Johnson LS, Kunsch CA;

XX MPI; 1998-272224/24.

XX N-PSDB; AAV27431.

XX Nucleic acid encoding antigenic peptide(s) from Streptococcus pneumoniae - or their epitope-containing fragments, useful in protective or therapeutic vaccines, and for diagnosis

XX Claim 11; Page 92-93; 118pp; English.

XX The present sequence represents a protein from Streptococcus pneumoniae. The nucleic acid sequence encoding the Streptococcus pneumoniae protein can be useful in vaccines for inducing protective antibodies against Streptococcus pneumoniae, for treatment or prevention of infection e.g. pneumonia, otitis media or meningitis. Probes based on the nucleic acid are used to detect Streptococcus infection (by usual hybridisation or amplification methods), also for isolating Streptococcus genes or their allelic variants. The protein can be used similarly to detect specific antibodies in standard immunoassays, especially for diagnosing or monitoring infections. Antibodies which bind the protein are used to detect corresponding antigens, to purify the protein and for passive immunisation (optionally coupled to a toxin). Vaccines are administered, e.g. by injection, orally or through the skin, typically at 0.01-1000 (especially 10-300) mu g/ml per dose.

XX Sequence 1007 AA;

Query Match 94.8%; Score 4171; DB 19; Length 1007;
Best Local Similarity 100.0%; Pred. No. 7e-72;
Matches 627; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 38 ETSODFEKKTAVIKEKEVSKNPVIDNNTSNEAKIKEENSKSGDYDTSFVNNTEN 97
    |||||
Db 1 ETSODFEKKTAVIKEKEVSKNPVIDNNTSNEAKIKEENSKSGDYDTSFVNNTEN 60
Qy 98 PKKEDKVVYIAEFKDESEGEKAIKEISLTKNTKVLTYTDRIFNGSAIETTPDMLDKIKOI 157
    |||||
Db 61 PKKEDKVVYIAEFKDESEGEKAIKEISLTKNTKVLTYTDRIFNGSAIETTPDMLDKIKOI 120
Qy 158 EGISSEVERAKVQPMNNHARKIEGVEAIDYLSKINAPFGKNPDGGMVSNIDTGDYR 217
    |||||
Db 121 EGISSEVERAKVQPMNNHARKIEGVEAIDYLSKINAPFGKNPDGGMVSNIDTGDYR 180
Qy 218 HKAMRIDDDAKASMRFKEDLKGTDKNYMLSDKI PHAFVYNGKITVEKYDDGRDYDP 277
    |||||
Db 181 HKAMRIDDDAKASMRFKEDLKGTDKNYMLSDKI PHAFVYNGKITVEKYDDGRDYDP 240
Qy 278 HGMHLAGIILAGNDTEODIYNFNGIDGIAIPNAQIFSYKMTSDAGSGFAGDETFHAIEDSI 337
    |||||
Db 241 HGMHLAGIILAGNDTEODIYNFNGIDGIAIPNAQIFSYKMTSDAGSGFAGDETFHAIEDSI 300
Qy 338 KKNVNVSVSSGFTGTGLVGEKYMQAIRALRKAGIPMVVATGNYATSASSSSMDLVANNH 397
    |||||
Db 301 KKNVNVSVSSGFTGTGLVGEKYMQAIRALRKAGIPMVVATGNYATSASSSSMDLVANNH 360
Qy 398 LKMTDTGNVTRTAHEDAIIVASAKNQTFEPDKVNIIGSESFYRNIGAFEDKSKITTNED 457
    |||||
Db 361 LKMTDTGNVTRTAHEDAIIVASAKNQTFEPDKVNIIGSESFYRNIGAFEDKSKITTNED 420
Qy 458 GTKAPSKLKFVYIGKQDODLIGDLRGKIAVMDRIYTKDLKNAFKAMDKGARAIVNVN 517
    |||||
Db 421 GTKAPSKLKFVYIGKQDODLIGDLRGKIAVMDRIYTKDLKNAFKAMDKGARAIVNVN 480
Qy 518 TVNYVNRDNTLPAWGVEADEGTSQVFSISGDDGVKLMNMINPDKTEVRANKNEDFK 577
    |||||
Db 481 TVNYVNRDNTLPAWGVEADEGTSQVFSISGDDGVKLMNMINPDKTEVRANKNEDFK 540
Qy 578 DKLEQYYPIDMESFNSNKENVGEKEIDPFKAPDTEKELYEKEDIIIVPAGSTSGPRIDLL 637
    |||||
Db 541 DKLEQYYPIDMESFNSNKENVGEKEIDPFKAPDTEKELYEKEDIIIVPAGSTSGPRIDLL 600
Qy 638 LKPDVASPGKNIKSTLNVINGKSTYGY 664
    |||||
Db 601 LKPDVASPGKNIKSTLNVINGKSTYGY 627

```

RESULT 5
ABP54664
ID ABP54664 standard; Protein; 1007 AA.

XX ABP54664;

XX 04-SEP-2002 (first entry)

XX S. pneumoniae SPI22 protein sequence SEQ ID NO:216.

XX Streptococcus pneumoniae; epitope; vaccine; antigenic protein;
KW antibacterial; Streptococcal infection; detection.

XX Streptococcus pneumoniae.

XX US2002061545-A1.

XX 23-MAY-2002.

XX 22-JAN-2001; 2001US-0765272.

XX 30-OCT-1997; 97US-0961083.

XX (CHOI/) CHOI G H.
XX (KUNS/) KUNSCH C A.
XX (BARA/) BARASH S C.
XX (DILL/) DILLON P J.

PA (DOUG/) DOUGHERTY B.
 PA (FANN/) FANNON M R.
 PA (ROSE/) ROSEN C A.
 PI Choi GH, Kunsch CA, Barash SC, Dillon PJ, Dougherty B, Fannon MR,
 PI Rosen CH;
 DR WPI; 2002-479261/51.
 DR N-PSDB; ABQ84899.
 XX
 XX New Streptococcus pneumoniae antigens, useful for detecting
 PT Streptococcus and for preventing or attenuating disease caused by
 PT Streptococcus infection -
 PS Claim 11; Page 50; 70pp; English.
 CC
 CC ABQ84792 to ABQ84904 represents nucleic acids which encode the
 CC Streptococcus pneumoniae antigens given in ABP54557 to ABP54669.
 CC The S. pneumoniae antigens have antibacterial activity and can be
 CC used in vaccines. The S. pneumoniae antigens can also be used to
 CC prevent or attenuate a Streptococcal infection in an animal. The
 CC polynucleotides encoding the S. pneumoniae antigens can be used to
 CC detect Streptococcus nucleic acids. ABQ84905 to ABQ85130 represent
 CC primers used in the cloning of S. pneumoniae ORFs (open reading frames)
 CC which are used in an example from the present invention.
 CC
 XX Sequence 1007 AA;
 SQ
 Query March 94.8%; Score 4171; DB 23; Length 1007;
 Best Local Similarity 100.0%; Pred. No. 7e-72;
 Matches 627; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 38 ETSODPFEKKTAVIKEKEVSKNPFVINDNTSNEAKIKEENSNSQGDYTDSPVKNKTEN 97
 DB 1 ETSODPFEKKTAVIKEKEVSKNPFVINDNTSNEAKIKEENSNSQGDYTDSPVKNKTEN 60
 QY 98 PKKEDKVVYIAEFKDSGEKAIKELSLKNTKVLTYTDRIENGSAIETTPDNLDKIKOI 157
 DB 61 PKKEDKVVYIAEFKDSGEKAIKELSLKNTKVLTYTDRIENGSAIETTPDNLDKIKOI 120
 QY 158 EGISSVRAQVQPMNMHARKEIGVEALDYLSKINAPFGKNPGRGKVTISNIDTGTDR 217
 DB 121 EGISSVRAQVQPMNMHARKEIGVEALDYLSKINAPFGKNPGRGKVTISNIDTGTDR 180
 QY 218 HKAMRIDDDAKASRFKKEDKGTDKNYMLSDKI.PHAFNYNGGKITVEKXDDGRDYFDP 277
 DB 181 HKAMRIDDDAKASRFKKEDKGTDKNYMLSDKI.PHAFNYNGGKITVEKXDDGRDYFDP 240
 QY 278 HGMHIAIGLAGNDTEODIKNFNGIDGIA.PNAQIFSYKNVSDAGSGFAGDETMFAIEDSI 337
 DB 241 HGMHIAIGLAGNDTEODIKNFNGIDGIA.PNAQIFSYKNVSDAGSGFAGDETMFAIEDSI 300
 QY 338 KANNDVSVSSGFGTGLVEKTYQAIRLRKAGIPMVVATGNTATSASSSSMDLVANNH 397
 DB 301 KANNDVSVSSGFGTGLVEKTYQAIRLRKAGIPMVVATGNTATSASSSSMDLVANNH 360
 QY 398 LKMTDGNMVRTAHEDAIAVASAKQTVBERDKYNIGSESPKYNTICAFPDKSKITTNE 457
 DB 361 LKMTDGNMVRTAHEDAIAVASAKQTVBERDKYNIGSESPKYNTICAFPDKSKITTNE 420
 QY 458 GTKAPSLKLFVYIGKGDODLIGDLRGKIAVMMDRIYTKDLKNAFKTAMDKGARAIVVN 517
 DB 421 GTKAPSLKLFVYIGKGDODLIGDLRGKIAVMMDRIYTKDLKNAFKTAMDKGARAIVVN 480
 QY 518 TVNYNRDNTLPLAMGYEADGEGTSQVFSISGDDGYKLANMINPDKTEYKRNKEDFK 577
 DB 481 TVNYNRDNTLPLAMGYEADGEGTSQVFSISGDDGYKLANMINPDKTEYKRNKEDFK 540
 QY 578 DKLEQYYPIMESFNSNKPNGDEKEIDPKFAPPTDKELVKEDIIVAGSTSGPRIDDL 637
 DB 541 DKLEQYYPIMESFNSNKPNGDEKEIDPKFAPPTDKELVKEDIIVAGSTSGPRIDDL 600
 QY 638 LKPDVSAFGKNIKSTLNVINGKSTYGY 664

DB 601 LKPDVSAFGKNIKSTLNVINGKSTYGY 627
 RESULT 6
 ABB98574
 ID ABB98574 standard; protein; 2835 AA.
 XX
 AC ABB98574;
 XX
 DT 14-JAN-2003 (first entry)
 XX
 XX Dextran saccharase, DSRE.
 XX Dextran saccharase; enzyme; cytosolic; dermatological; antiseborrheic;
 KW DSR-E; glycosyl transferase; dextran; prebiotic; pharmaceutical;
 KW microflora regulation; intestinal transit; mineral assimilation;
 KW colon cancer; acne; dandruff; body odour.
 XX
 OS Leuconostoc mesenteroides NRRL B-1299.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..40
 FT /label= signal_peptide
 FT 41..2835
 FT /label= Mature_protein
 FT 423..439
 FT /label= SEQ_ID_6
 FT /note= "This sequence is specifically claimed in Claim 7"
 FT 478..501
 FT /label= SEQ_ID_7
 FT /note= "This sequence is specifically claimed in Claim 7"
 FT 519..539
 FT /label= SEQ_ID_8
 FT /note= "This sequence is specifically claimed in Claim 7"
 FT 560..571
 FT /label= SEQ_ID_9
 FT /note= "This sequence is specifically claimed in Claim 7"
 FT 631..645
 FT /label= SEQ_ID_10
 FT /note= "This sequence is specifically claimed in Claim 7"
 FT 1014..1021
 FT /label= SEQ_ID_11
 FT /note= "This sequence is specifically claimed in Claim 7"
 FT 2120..2138
 FT /label= SEQ_ID_12
 FT /note= "This sequence is specifically claimed in Claim 7"
 FT 2161..2184
 FT /label= SEQ_ID_13
 FT /note= "This sequence is specifically claimed in Claim 7"
 FT 2202..2214
 FT /label= SEQ_ID_14
 FT /note= "This sequence is specifically claimed in Claim 7"
 FT 2243..2250
 FT /label= SEQ_ID_15
 FT /note= "This sequence is specifically claimed in Claim 7"
 FT 2315..2332
 FT /label= SEQ_ID_16
 FT /note= "This sequence is specifically claimed in Claim 7"
 FT 2689..2696
 FT /label= SEQ_ID_17
 FT /note= "This sequence is specifically claimed in Claim 7"
 FT 1981..1112
 FT /note= "Catalytic domain, SEQ ID 1. This sequence is specifically claimed in Claim 4"
 XX
 PD 20-SEP-2002.
 XX
 PF 19-DEC-2001; 2001FR-0016495.
 XX
 PR 16-MAR-2001; 2001FR-0003631.

XX (CNRS) CNRS CENT NAT RECH SCI.
 PA Boconnet SAM, Renaud SMMC, Willemot RML, Monsen PEF;
 PI WFL, 2002-715213/78.
 XX N-PSDB; AB080961, AB080962.
 DR
 XX New glycosyl transferase enzymes, containing glucan bonding and
 PT catalytic domains and producing alpha-(1-2) branched dextrans, useful
 PT in probiotic, pharmaceutical or cosmetic compositions -
 XX Claim 6, Page 65-74, 82pp; French.
 XX
 XX The present sequence is a novel dextran saccharase, DSR-E, from
 CC Leuconostoc mesenteroides NRRL B-1299. The saccharase has glycosyl
 CC transferase activity suitable for producing dextrans having alpha(1-2)
 CC branches from sucrose, p-nitrophenyl-alpha-D-glucopyranoside,
 CC alpha-fluoroglucose, alpha-D-glucopyranoside-alpha-D-sorbiopyranoside or
 CC alpha-D-galactopyranosyl-sucrose. The dextran saccharase is useful in
 CC prebiotic, pharmaceutical or cosmetic compositions. The dextrans and
 CC related compounds having alpha(1-2) bonds, produced using DSR-E, may be
 CC involved in signalling/cellular recognition processes in vivo
 CC (specifically in regulation of microflora in the intestines or on the
 CC skin); and are potentially useful for improving intestinal transit,
 CC increasing assimilation of minerals (e.g. calcium and/or magnesium),
 CC preventing cancer of the colon and combating skin problems such as acne,
 CC dandruff and body odour.
 CC
 XX Sequence 2835 AA;
 SQ
 Query Match 35.2%; Score 1549; DB 23; Length 2835;
 Best Local Similarity 24.7%; Pred. No. 5.2e-16;
 Matches 289; Conservative 150; Mismatches 149; Indels 581; Gaps 219;
 QY 1 YPVVLA-DTSSSEDALNIS-D--KEKV---AE-NKEKA---E---N---I--HSM 37
 DB 524 Y-RVAVNVDV-DLLOIAGDYFKAAYGTGTEANNNHISILEDNDNSAIVAKHGN 581
 QY 38 ETSODFKE--K-KTAVIKEKEVY-S--KNPVIDNNTSNEAK-IKENSKSGQDYT--- 87
 DB 582 QUTMPFPAHLAKVIALNM-PLAOSGLE-PLI--NTS--LVKRGKATENAPNAPFR 635
 QY 88 --DS-----FV-----NK-NTE-----NPKEDKV-----VYIA-EFK-DKE-SGEKAIX 121
 DB 636 AHDSEVQVYIAQIINDKINTKSDGLTVP---DEIQATIVNADLKADEKVT-AVNP 691
 QY 122 -ELSS-LKN--TKV--LYTY-DRIF--NG---SALETPDNLDKIQIGISSVERAQK- 168
 DB 692 ASYAVLITNKOT-VPRVY-YGD-LFSDDQYMSQ-K-SPY-YDAITSL--LKS--RI-KY 740
 QY 169 V--OPM-MN--H-----ARKEI---GVEAIDYIK-SINA-PFGKNPGR---GMVVS 208
 DB 741 VAGGSMMNTYIHECPDPAKMETKPGVLTSVRYGAMTADDLG-NSDTRQOGIGLVIN 799
 QY 209 ----NI--D-----TGT-----DVR--H--KA-MRI-DDDAKA-----S--M-R--FK 234
 DB 800 NKPFLINDDEQIVLNMGAHKNQAVRPLMLTTKSLQIYDDAGA.PVYYTNDAGQLIF- 858
 QY 225 KEDL-KGTDKA'-----Y--WL-----SD-----K--IPH--A----- 254
 DB 859 KSDMYGYV-SNPQVSGYFPAWVPVGSQDARTOSSQSEKDGVDYHSMALDSNVIYE 917
 QY 255 -F-NY-----Y-----NG-----GKIT--VE--KY-----D-----G-- 271
 DB 918 GFSNFQAMPEKDDFTNVKIAQNAKLFKDLG-ITSELAPQYRSSSTDSFLOSVIONGYA 976
 QY 272 ---R-D--YFDB--HG-----MHAIAGILAGND--TEOD--IKNFG--I--DSIAN 307
 DB 977 FTDRDYGVNPTKKGTVUQGLDSLRAHAGIQAINDWVP-DQIYNLGEQIVTAVRN 1035
 QY 308 AQI-PSYK-----MYS-----AGSG-----FAG-----DE-----TMFH--AIEDSI-KIN 340

DB 1036 GSGKDYDYSVINNTLY-DSRTVGGGEYQEKFGGLFLDQLKQVPSLETRKQI-STNOPMN 1093
 QY 341 VDVVSV---SSG-FTGT---GLVGEKY-----W---Q--AIRALRKAG-IP--WV---VAT 378
 DB 1094 PD-VKIKEMSAKKFNSNIOG-RGAMYVLKDMATNQYFNVS--DNGFLRKQLLGKRTST 1149
 QY 379 G-----N-----YATS---A--S-----SSWDLVANNHLKXTDTG--NVTPTAHED--- 414
 DB 1150 GFITTEGKTSFYSTSGVQAKDPTFIQGTWNYVFDNAGY-WL-TGKQNI-----H-DKNY 1201
 QY 415 ----ALAVASA---KNQVPEPKVNIIGSEFK---Y--RINGA-FEDKS-K-I--T-- 453
 DB 1202 FLPNQVLDYALFDNQ--EF--YNNKAGEQVNNQYQDSQONQWYFFENGRAALIGLTV 1258
 QY 454 TNEDETKAPSKLKF---VYI-G---KGQ-DQ---D-----LI---G-L-D-----L- 483
 DB 1259 PNADGTHV-TQY-FDANGQIKGTALKDQNNQRLRYDEALGNVNVSWGLADKSWLYLN 1316
 QY 484 -RGKIAVM-DR-IYTKDLKNA--FKKAMD---KGARAIMVNTVN---YNNRD----- 525
 DB 1317 AQG-VAVTGNQKI---DGE-EYVF-NA-DGKQYKGN-AI-IDN--NGDQRYDQDKGVMV 1365
 QY 526 -N-WTELLPAMG---YEADG---T-----KSQV-FSISGDDG--VK-----L---W--- 557
 DB 1366 VNSWGLPD-GSWLYLNDKGIAYTGROVINNVNF--FGNDKQIKDAFKLLSDGSWYL 1422
 QY 558 -NM-----INP-----DK-----K-----TEV--KR-----N-----NK--E--D- 575
 DB 1423 DDKGLITTGAKVINGAMFPDQKHQIKGDASTIDANGKHYYDKNDCHLVNWSGELPDG 1482
 QY 576 ---F--K--DKL-EQYYPIDMS--F-----NSNKP--NVGDEKEI---DFKEAPD 611
 DB 1483 SMLYLEEGQDAVYQGRV-IDGKTRYFDEDEGQIKNSLKTIAN-GD-K-IYLDGD-GVA-A 1536
 QY 612 T-----DKELY-KED---I---VPA--GSTW-----GP-RID---LLKPD 641
 DB 1537 TGLQHVGDKIMYFDEDEGQVKGFSVAKDG--SWYLLNDGVAAGVSSYINGOSLYPDQ 1594
 QY 642 VSAPGKNIKSTLN-VIN--GKST--Y-GY 664
 DB 1595 ----GQVKA--YNEVNSDG-TTNYTTGL 1616
 RESULT 7
 AAR08221
 ID AAR08221 standard; protein; 1475 AA.
 XX
 AC AAR08221;
 XX
 DT 25-MAR-2003 (updated)
 DT 06-MAR-1991 (first entry)
 XX
 DE Recombinant alpha amylase pullulanase enzyme.
 XX
 KW Starch; pullulan; saccharification; debranching; liquefaction;
 KW hydrolysis.
 OS Clostridium thermohydrosulphuricum DSM 3783.
 XX
 XX Key Location/Qualifiers
 FH Peptide 1..31
 FT /label= signal peptide
 XX
 PN EP402092-A.
 XX
 PD 12-DEC-1990.
 XX
 PD 05-JUN-1990; 90BP-0306088.
 XX
 PR 05-JUN-1989; 89US-0361368.
 XX
 PA (ALKO-) ALKO LTD.
 PA (ARCO-) ARCO OY AB.

Host cell containing recombinant expression vector encoding Clostridium botulinum type B or E toxin - useful to treat humans and other animals at risk of intoxication with clostridial toxin

Example 15; Page 220-230; 428pp; English.

This is the amino acid sequence of Clostridium difficile toxin A, deduced from the coding region (see AA050560) of the toxin A gene. Toxin A is a potent cytotoxin that plays a direct role in damaging gastrointestinal tissues. Severe cases of C. difficile intoxication result in pseudomembranous colitis. This would be prevented by neutralising the effects of toxin A in the gastrointestinal tract. Examples are provided of the production of recombinant C. difficile toxin A in host cells and of the in vivo neutralisation of toxin A by antibodies against recombinant toxin A polypeptides. The invention specifically relates to recombinant proteins derived from Clostridium botulinum toxins (see AA068369-400) and their use as immunogens for the production of vaccines and antitoxins.

Sequence 2710 AA;

Query Match 34.7%; Score 1525; DB 19; Length 2710;
Best Local Similarity 23.8%; Pred. No. 1,3e-15;
Matches 300; Conservative 165; Mismatches 137; Indels 659; Gaps 258;

1 YPVVLA-D---TS--S-S-E-----D---ALNISDKE---K---VAENKEK---H 30
1426 YSLLSDSKVYLNLSNTEIKNTLGLDSKNIAVNTD-ESNNKYGAISKTSQKSIH 1484
31 ----ENI---H---SAME-TSQDF-KEKTAIVIKEV--VS-KNPVIDNNT----- 67
1485 YKXDSKNILEPYNSTLEFNSKDFIAE-DIVFMKDDINTITGK-YVVDNNTDKSIDPSI 1542
68 ----S-S-NEAKIKENS---NKS-OGDYTDSFY-NK---NTENKPK--EDKV-VY-I-A- 108
1543 SLVSKN-QVKV---NGPLYNESVYSYLD-FVKNSDGHNTSMFNLFDNISFWKLGFG 1597
109 E---F-KDKE-S--GEKA---IKELSLKVTK-V-LY-----T-YDR--IF--NG--SA 143
1598 ENINFDVDKFTLVG-KTNLGYV-EF-ICDNNKIIDYFGEMKTSKSKTIFSGANGRV 1654
144 IET--TPDN-LDKIKO-I---E---GISSEYER-AQRY-OP-----MMNRAREKIVE- 183
1655 VEPIYNDTGED-ISTSLDSFYEBLYGI---DRYINVLAPLDYLTSLIN-----INTNY 1705
184 EALDYLSI---NAPFGKN-FDGRGV-ISNID-----T-GTD-----Y----- 216
1706 YSNEYEPIIVLN-P---NTF--HKQVNI-NLDSSEPEYKMTSGSDFILVRYLESNKK 1758
217 ----RIK-----AMRID-DPAKA-S---MR-FK---K-E---D-LKG---TD 242
1759 ILQKIRIKGLISNQSFRKSIDPKDIKLSLGYTMSNFSFNSENELDRDHL-GFKIID 1817
243 -KNYML-SD---K---I-----PHAFNY--Y---NGK---I-T-VB--KYD- 269
1818 NKTYVDEDSGLVGLININNSLFYDPPIEBNLVTGMQTNGKKYVFDINTGALTSYKI 1877
270 -DGRD-YF--DPHG-MHAGIAGND-----T-----E-ODI---KNF---NG-- 300
1878 INGHGFYFND--GVML-GVFGKPDGEYFAPANTNNIEGAIYVQSK-FLLNGKK 1933
301 --ID-----G---I-----APN---A---Q-I--PSYKNV--SD-A---G--S- 321
1934 YFPDNNKAVTGMRIINNEKYYFNPNNALAVGLQVINDNKT--YFNPDTALISKGMQTV 1991
322 -G---FAGD-ET--MF--H-AIEDS1KH--NYD-VVSVSSG-F-TGTG--LVGEK--Y 360
1992 NGSRYF--DTDTALAFNGYKTI-DG-KHFFYDSDCVYKI--GVFSISNGEYFAPANTY 2045
361 ----WQAIRLR-K---AGIPNVV-----A-TG---N---Y-AT-SA-SSGSW-- 390
2046 NNNIEGAI-VYQSKFPLTNG-KKYVPDNNKAVTGLQITDSKKYVFNNTVAEAATMQT 2103

QY 391 -D---LVAN-NHUKM-----T-DTG-----NTRTA-A-----H-----ED---AI 416
DB 2104 IDGKYYFNTNTAEATGMQITD-GKKYYFN-INTALASGYTILINKHFFYNTDGMQI 2161
QY 417 AVASAKQVTE-F-----DKVNIGGESFKYRN-----I-GA--FF--DKSK-IT-----TN 455
DB 2162 GVFKGPN-GFEYFAPANTDANNIEGAILYQNEFLTLNGKKYFGSD-SKAVTGMRIINN 2219
QY 456 EDGFK--AP-SLKPV-Y---IKGQDQDILGD--LR-GKIIVMDR--IYTD----- 497
DB 2220 K---KTYFNNALIAATHLCTINN--DKYFSYDGLQNGYITI-ERNFY-EDANNESK 2272
QY 498 ----LK--NAFK---KA--MD---KGARIMV-N---TVN---YY--NRD-----NW- 527
DB 2273 MVTGVFGRPGFEPYAPANTHNNNIS-QAI-VYQKFLTLNKKYFEDN-DSKAVTGMQ 2339
QY 528 T---E---L---PAM-GYEA-DEG-----T-K-S-Q-V-----F--S-I-S 549
DB 2330 TIDGKYYFNLNTAEATGMQITD-GKKYYFNLNTAEATGMQITD-GKKYYFNTNFIAS 2388
QY 550 -G--D-DGVK-LM-N---M-I-----N-----P---DKT-E--V-KRN-----N- 572
DB 2389 TGYTSING-KHFFYNTDGMQIGVFKGPNGEYFAPANTDANNIEGAILYQNKFLTLNG 2447
QY 573 KE-DF-KD-K---L-----EQY-----PID-ME-SFNSN---KP--NV- 598
DB 2448 KKYVFGSDSAVVGRLTIDGKYYFNTNTAVATGMQITNGKKYFNTNISTASTYTI 2507
QY 599 -G-----DE-KEI-----D-FK-PAP-DTK--E-----LYKEDIYVAG- 626
DB 2508 SGKHFFYNTDGMQIGVFKGPDGEYFAPANTDANNIEGAILYQNFYVFGN 2567
QY 627 -S---TSWGRID---LLKPDVSAPGN-IKSTL--N-VI-NG-----K-STG-- 663
DB 2568 NSKAATGWT-IDGNRYFEPNTAM-GANGYK-TIDKNYFPRNGLPQIGVFGSN-GFE 2623
QY 664 Y 664
DB 2624 Y 2624

RESULT 10
ABB47329
ID ABB47329 standard; protein; 1530 AA.
AC ABB47329;
XX 05-FEB-2002 (first entry)
DE Listeria monocytogenes protein #33.
XX
KW Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
KM vitamin B12; bacterial infection; disease.
OS Listeria monocytogenes.
XX
PN NO200177335-A2.
XX
PD 18-OCT-2001.
XX
PF 11-APR-2001; 2001MO-FR01118.
XX
PR 11-APR-2000; 2000FR-0004629.
XX
PA (INSP) INST PASTEUR.
PI Buchrieser C, Frangeul L, Couve E, Ruanjok C, Peith H, Dehoux P;
PI Duesenget O, Chetoui F, Nedjari H, Glaser P, Kunet P, Cosserat P;
PI Daniels U, Goebel W, Krefic U, Kuhn M, Ng B, Vazquez-Boland JA;
PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;
PI Chakraborty T, Dommann E, Hain T, Berche P, Charbit A, Durant L;
PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;

P1 Maduenio E, De Pablos B, Wehland J, Kaerst U, Entian K, Hauf J;
 P1 Rose M, Voss H;
 XX WPI; 2002-010914/01.
 XX
 PT Genomic sequence for *Listeria monocytogenes*, useful e.g. for treatment
 PT and prevention of *Listeria* and related bacterial infections, and
 PT related polypeptides -
 XX
 PS Claim 6; SEQ ID No 34; 192pp; French.
 XX
 CC The present invention relates to the genome sequence of *Listeria*
 CC *monocytogenes* EGD-e (see ABA03041). The genome sequence and fragments of
 CC it are useful for selecting probes and primers for detecting genes in *L.*
 CC *monocytogenes* and related organisms, and for studying genetic
 CC polymorphisms and other genomes. The present sequence is a protein
 CC encoded by the genome sequence of the present invention. Proteins
 CC expressed from the genome sequence are useful for raising specific
 CC antibodies, identification of *L. monocytogenes* and related organisms, and
 CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin
 CC B12. The genome sequence and protein encoded by it are also useful for
 CC selecting compounds that regulate gene expression and cell replication
 CC and modulate *L. monocytogenes*-related diseases. In addition, the genome
 CC sequence and proteins encoded by it are useful in pharmaceutical and
 CC vaccines compositions for the treatment or prevention of infections by *L.*
 CC *monocytogenes* and related organisms.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIRO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX
 SQ Sequence 1530 AA;
 Query Match 34.4%; Score 1511; DB 23; Length 1530;
 Best Local Similarity 24.6%; Pred. No. 1.9e-16;
 Matches 302; Conservative 162; Mismatches 127; Indels 635; Gaps 248;
 QY 1 YPV-----V-LA-DTSSS-E---DAL-N--ISDKE-KVAENKE--KHENHSAME- 38
 DB 163 FPISENDESEFTIKKIPDTNNAFEKKSGSTVQNAQIWEIKI--NQLATHENA-SV-ED 218
 QY 39 ---TSQ---D-FKE--KKTAVIK-EKEVSK-NP---V--IDNNTNEAKKE--EN 78
 DB 219 FPGNQOKILPFSF-EVHEMNMHIDGTVEIGAKVDPSEYTVTKI--NNTSG-DTGFRLTREN 275
 QY 79 SNKS--OGDY---TDSFVNK---NTE--N--PKKE--D-KVY-----YIAE-----EKD 112
 DB 276 PIDSAYQITTYQTEITD-F-NKATTSNATFNSDETEEVMSDQVLPDPPIIKQKPLFL-D 332
 QY 113 KESGEKAI--K-ELS---SLKNTKVLV-TY-D-----RIFNG-SAI-----E-T-T 147
 DB 333 TTSN-K-ISMKEVYNGHYALDNS-VMDTYIDEGSPGIGIRIVAGDGIPLLPTEYIVT 389
 QY 148 -----PDNID--KI--KQI--EG---ISSVERAKQVPM--NHAKKEIG-- 181
 DB 390 DMGSAGFKIEI-PD--DGYITTYFETIVEGETGMISNT--ASIIISPNDNEA--E-GDY 442
 QY 182 -VB-----EA-IDYIKSI-NAPF-----GK-----NFDGGMVSN--I-DTGTD--Y- 216
 DB 443 FIPDPSPLEGLID--KLIEN--FNAKTGLTWEIITIKNG-G-TLHNPVITDEPDDGLI 496
 QY 217 RHKA-MRI-DDD-AK-ASMRPKE--DKGTDKNTW-----LS-D-----K--I 251
 DB 497 FHPATLKIMDEBOVLDSADY--EVVPLNG--DSS--WKGQFQINFRDITGQHVTITKQV 552
 QY 252 -P--HAF--N-YVNGKI--TVB-----KY-D--DGRD-Y---FD-PHGMIAG-- 284
 DB 553 NPSTHSGDNEYNNATIKTDTAEASDSDTKWIDKVIDA--DGYKNGVFNKYTG--EIEWKL 610
 QY 285 ILANGD-----T-EODIKNF--G---I-DGI-----A--NNAQIFSF--KMSD 318
 DB 611 IF--NDSSKLSIKSPITE-D--SLNSGQTFIDDSIEHIKIDISATP--QVGEIIPENY-D 662
 QY 319 AGSGFA-----GDETW---F-----HAIE--DS---I-KH---N--V--D---VVS--- 345

DB 663 V-T-FTKKENGNEQMLTTFKKPLIIPVEVYTKTPVGTITPKYKNAVISDGEVLADYE 720
 QY 346 --V-----SSGETG-TGL-VEEK--YWOAIRALRKAG--IP-MVVA-----TG-NVATSA--- 385
 DB 721 AEVIDDNANKRVNKSQVQDNDIME--IVANQ--SGSTVSNATVDTIGTGKLTSSIKV 778
 QY 386 --S-S-SWDLVA--NNHAKMT--D--TG-----N-VTRT--AAHEDAI 416
 DB 779 YKSQTSVTKMKLOESN--MPSISPGYDLKTGVDEESNLEYQVFKKEINISYVAKYQAI 837
 QY 417 AVAS-AKN--Q---TVEP--DKVNI--GGESEFKRNI-----GAFD-----KS-KITTN 455
 DB 838 TLTSDETTLTQIGNSVFTGPD--NITGGEKTEKNIIEVKITTG--DGTGGEKIKILN 892
 QY 456 E-DGTRK-PS-KLK---F-VY-----I-----GK-G--Q--DQDIL-GUD--LR-----G 485
 DB 893 KVD--KADPSIPLEGATFDLYANDEKVDOTTDNKGVIKFD-DLYVG-DYTLKVSAPEG 948
 QY 486 -----K-IAVW---D---R-----I-----YTK-D-----LKNA-FKKAM- 506
 DB 949 YTLPTASTENIQVLEBQDERKVQVMEKMPKENGVEHLVTKDCAATLGAFAE--SLY 1006
 QY 507 DK-GA--RAIMVY--N---TV-N-----YNRDNWTLBPANGYADEGTS--QVFSI-SG 550
 DB 1007 DKSGAEIQLNGILTDENGELTIHNDLGSYYLKE--TRAPE-GYKLSE--KTWE-FSVESG 1060
 QY 551 -D-----D-G-----V-----KL-----NMINPD--K--KT-----EVK 569
 DB 1061 QVDAIEIOAENENKOLGAVLTKVDSETNALKSGAKFNILN--DSGEVYQTNILVSDENGELI- 1118
 QY 570 R-RNKE--D--FKDLERQ---Y-----YF--I-----DM--ESFSN--KPNV-- 598
 DB 1119 RVQLBLEGDYAFQ--ETLAPNTNYDATNTWPTIYAGQTSAMVYAE--N-NKTKDPDVT 1174
 QY 599 G-----D---EK-E--I--DFKAPD-----TDK--E-----L-----Y--KEDI 621
 DB 1175 GEVLIVKQDSATGTELTGAVFDMTWA--DGAIVASNLTTDANGELTYTNLAPGKYSFKS-T 1232
 QY 622 IVPAG---STS-W-----GP-R---IDL-----LKPVSAP--G--KNIK--STLN-V--- 655
 DB 1233 KAPGEYELADVWVEFTTAPNQPEKITTTAENTKLAP-I--PDAGSVKIKIKODESNGVRLA 1289
 QY 656 -----I--NGKS--T-Y-----G-Y 664
 DB 1290 GAERSLAENGETIQTNLKTDENGEL 1315
 RESULT 11
 AAM09255
 ID AAM09255 standard; Protein; 1938 AA.
 XX
 AC AAM09255;
 DT 27-MAR-1997 (first entry)
 XX
 DE Bacillus alkaline amylolupulinase.
 XX
 KM Amylolupulinase; alpha-amylase; pullulanase; detergent additive.
 XX
 OS Bacillus sp. KSM-Ap1378 (FERM BP-3048).
 XX
 FH Key
 FT Peptide
 FT 1..32 Location/Qualifiers
 FT /label= Sig_peptide
 FT 33..1938
 FT /label= Mat_protein
 FT /note= "mature amylopululanase"
 FT 33..833
 FT /label= Alpha-amylase
 FT /note= "N-terminal alpha-amylase"
 FT 462..645
 FT /note= "amyIase active centre"

Query Match	34.3%	Score 1509;	DB 17;	Length 1938;
Best Local Similarity	25.1%	Pred. No. 6.1e-16;		
Matches	292;	Conservative 166;	Mismatches 128;	Indels 578; Gaps 240

1 Y-PV-VLA--D-TSSS-E--D-AL-----ISDKE-KVAENKE-K-HENIHA 36

D	b	25	YLPAPAAAE	TDKDKRIE	SFYERPD	GNBYG	MLWMG	TGVK	QDQID	PTFEB	KEG	KA	YAD	I	-A	82
Q	y	37	METSD	QF	EXK	ET	A	-VI	-K	----	E	K	E	V	----	77
D	b	83	--VS-	D	-NAD	V	G	F	I	R	K	G	D	M	E	135
Q	y	78	NSNK	-S	O	G	D	T	S	F	V	-N	K	T	E	122
P	b	136	GSP	P	V	A	D	G	N	-A	D	F	F	R	D	189
Q	y	123	-L	S	L	K	N	-K	T	L	N	Y	----	D	-	150
D	b	130	T	L	S	D	P	-G	E	E	E	Y	L	T	V	248
Q	y	151	L	D	K	-I	K	-Q	I	E	S	S	V	E	R	196
D	b	249	Q	N	A	V	K	D	I	E	----	S	E	E	-T	300
Q	y	157	G	-K	N	F	D	G	K	W	I	S	N	I	D	234
D	b	301	G	V	K	N	L	P	----	I	T	A	I	D	E	353
Q	y	235	K	E	D	L	K	G	D	-K	N	Y	M	S	D	271
D	b	354	--D-	G	-D	S	S	N	----	N	D	-P	H	G	I	402
Q	y	272	----	R	----	D	-Y	F	----	D	----	P	H	-G	-M	289
D	b	403	V	D	N	I	K	P	V	R	S	E	G	P	D	461
Q	y	290	D	T	E	O	D	I	K	----	F	N	G	I	-P	328
D	b	462	D	V	----	V	L	N	H	T	G	-G	L	K	P	513
Q	y	329	M	F	H	A	I	E	-D	S	I	K	A	N	Y	369
D	b	514	-F	L	T	-E	N	P	D	-R	E	O	V	W	O	569
Q	y	370	A	G	I	P	-M	V	-A	T	-G	-N	----	Y	A	400
D	b	570	A	-M	E	H	K	I	G	E	A	M	A	N	V	627
Q	y	401	-T	D	T	G	N	-V	T	-R	-T	A	A	H	-E	428
D	b	628	N	S	K	A	N	T	A	T	T	L	Q	P	L	687
Q	y	429	-D	-K	V	----	N	----	I	-G	S	E	F	K	-Y	464
D	b	688	N	D	V	P	Y	T	R	O	M	P	D	D	V	743
Q	y	465	L	K	F	----	V	I	G	----	K	G	O	D	L	503
D	b	744	L	L	F	S	R	T	G	E	N	S	V	Y	G	800
Q	y	504	K	A	M	-D	K	G	A	R	A	I	M	-V	-N	527
D	b	801	P	A	M	E	D	-G	-T	L	L	E	N	A	V	858
Q	y	528	--T	E	----	L	P	-A	M	G	----	Y	E	A	-D	556
D	b	859	A	A	P	S	E	N	M	P	G	S	G	T	P	918
Q	y	557	----	N	M	----	I	----	N	P	-D	-K	T	-E	-V	594
D	b	919	P	O	L	N	E	I	W	I	K	O	G	S	D	973
Q	y	595	K	-P	N	----	V	G	----	D	-E	-K	E	I	D	617
D	b	974	G	M	P	G	A	D	A	A	I	G	K	Y	A	1033
Q	y	618	K	-E	D	-I	----	I	-V	-A	-G	----	S	T	S	647
D	b	1033	K	E	G	D	K	V	T	N	E	G	V	A	P	1088

[illegible]

QY	123	LSLSKNT-KYLTYK----	D-R--I-----FN--G-SAIE--TT-----	P---D-N	150	
Db	190	TLSDLP-TGEHRYTLVTDGTE	EEVDPYNTVDRSRVVEYTSVQV	SASFTIPAKVDYN	248	
QY	151	LDK-1K-OIEGSSVERAKQV	QPM-MNHARKBI-GVEEA-ID-Y-L-----	K-SINAPF	196	
Db	249	QNAVVKVDIE--SETE--TKI	REMSINLS--EIGSEKATIDP	ALNELTVAVKQVTA--	300	
QY	197	G-KNFDGGMWISND-TGTDY	RHK--A-----NR-I-----	D-D-DAKA--SM--RPF	234	
Db	301	GVKNLP-----ITALDERGN--	RHGSATLEVOARTITGEKAD	PFMDOSVVFVMTLDRFF	353	
QY	235	KEDKGTG-KNYWLSDKI	PHAFNY-----YNGS--K-ITVEKD--D-G-----		271	
Db	354	--D-G-DSSN--ND--PHG	IGDTSKSGTQSGDPFGIT--QR	LDYDELGINTIWSP	402	
QY	272	-----R-----D-YF-----	D-----PH-G-M--H--I--A--GI-LAGN		289	
Db	403	VVDNIKFDVRHSEGPDT	PYAVYHGYMAWNPGBLNP	HPGSMADFHEMIDAAHERGJIKIM-V	461	
QY	230	DTEODIKN--FNGI--D-GIA--	PN-----AQFSIKYTS	SDAGS-----FAG--DET	328	
Db	462	DV--VIAHRTG-GIKP	GDSSSVANFPDEDRAR-FD-GMLRDGSG	GEVRGELALPD--	513	
QY	329	MEHAIE--DSIKHNV-----	D-V-VS-VSSG-----F-TGT-GLVGEK-Y	WQAIR-ALRK	369	
Db	514	-FLT-ENND-VREQVQWQ	WTMIEKSRFAKNTIDYF	AVDVIKAV-EDTTMAF	FNALTK	569
QY	370	AGIP--MVV-AT-G--N-----	YATS--ASS--SSW--DLVANHL-----	KM--	400	
Db	570	A-MPEHKLIGBAMGANV	ADDGLYLSNGMWSLLD	DFPKYARDF-ANGQLDAVQCKEAR	627	
QY	401	-TDIGN-VT--R-TAAH-ED--	AI-----AV-AS--AKNQVY-----	EE--	428	
Db	628	NSKLNNTATLQOFLGSH	EDRPFYEVEBGDLKRYO	VASLQLTAKQPVITYYGEELGLPGK	687	
QY	429	-D-KV-----N-----I-G	GESFK-YRNIGAF-----F--DKSKITTEDGT	KAPSK	464	
Db	688	NDPYTYTRQWMPMDVD	GNELBHYQKTLAFRNDNP	PTFAKGRKKYA-----GSDSGY	743	
QY	465	LKF-----VYIG--KQD	ODLIGDLRG-KIAVMRIYT-KDL-KN-----	AFK-503		
Db	744	LLFSRTYGENSVYVGL	NTAAAKD-VTLNF-GSSEAV	VTDR-YSGOEYQANEEGQVTFESI	800	
QY	504	KAM-DXGARAIMV-V-N-----	T-----V-N-Y--YNR-DN-----	W--	527	
Db	801	PAMED-GG-TVLL	EENAVPPEBEPEPESIBENTIR	IRIHYORDNSYENLGIMLMDV	858	
QY	528	--TE-----LP-AMG--YEA--	D-EGTK-SQ-V-FSI--S-GD-DG--VKL--		556	
Db	859	AAPSNNPSSGTPFGAG	VNTDYGAIVVDVLEADQNI	CFVLANTNGDKDGGDKAVELFS	918	
QY	557	-----NNM-----I--NP	D--KKT-EV--KRNKED	DFD-KLEQY-YPIDMS	FPNSN	594
Db	919	PDLNEIWIKQSD	DEVFLYPEVDLPANTVRIHYERN	-ADYEGWGL--NNWR-DVES-ESD		973
QY	595	K-PN-----VG-----D-E	KEID-----FKPA-P-----D-T-D--K--EUY-		617	
Db	974	GMPGADADAAGIGK	GAYDIKLB-DANKIGFLV	NKQSGQGTDMTFMLKQYNQLPV	1033	
QY	618	K-ED-I-----IVP-A--G-----	STSMGR--IDL-L--LK-PDVS	APQK	647	
Db	1033	KEGEDKYVNTPYGTV	PLALVSGEVLSDKILSLTF-TRTEG	DLBELKQLLEIKVD--G-	1088	
QY	648	N-IKS-TIANV-INGKSTY--	--G-Y	664		
Db	1089	NDV-SFT-DVTIE	EKTVHVGEP	1110		

XX ABP30377;
AC 02-JUL-2002 (first entry)
XX Streptococcus polypeptide SEQ ID NO 9930.
XX
DE Streptococcus polypeptide SEQ ID NO 9930.
XX
KM Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KM group A streptococcus; Streptococcus pyogenes; antibacterial;
KM antiinflammatory; infection; vaccine; meningitis; gene therapy.
XX
OS Streptococcus agalactiae.
XX
PN W0200234771-A2.
XX
PD 02-MAY-2002.
XX
PF 29-OCT-2001; 2001WO-GB04789.
XX
PR 27-OCT-2000; 2000GB-0026333.
PR 24-NOV-2000; 2000GB-0028727.
PR 07-MAR-2001; 2001GB-0005640.
XX
PA (CHIR-) CHIRON SPA.
PA (GENO-) INST GENOMIC RES.
PI Telford J, Maignani V, Margarit Ros YI, Grandi G, Fraser C;
PI Tettein H;
XX
DR WPI: 2002-352536/38.
DR N-PSDB; ABN71008.
XX
PT New Streptococcus protein for the treatment or prevention of infection
PT or disease caused by Streptococcus bacteria, such as meningitis, and
PT for detecting a compound that binds to the protein -
XX
PS Claim 1; Page 4122; 4525pp; English.
XX
XX The invention relates to a protein (ABP25413-ABP30895) from group B
CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
CC the specification. The proteins have antibacterial and antiinflammatory
CC activity. (1), nucleic acids encoding (1), ABN6044-ABN71526 and
CC antibodies that bind (1) are used in the manufacture of medicaments for
CC the treatment or prevention of infection or disease caused by
CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
CC Nucleic acids encoding (1) are used to detect Streptococcus in a
CC biological sample. (1) is used to determine whether a compound binds to
CC (1). A composition comprising (1) or a nucleic acid encoding (1), may be
CC used as a vaccine or diagnostic composition. The disease caused by
CC Streptococcus that is prevented or treated may be meningitis. Nucleic
CC acid encoding (1) may be used to recombinantly produce (1) and may be
CC used in gene therapy. Antibodies to (1) are used for affinity
CC chromatography, immunoassays, and distinguishing/identifying
CC Streptococcus proteins.
XX
SQ Sequence 2274 AA;
Query March 34.2%; Score 1506; DB 23; Length 2274;
Best Local Similarity 25.6%; Pred. No. 1.5e-15;
Matches 280; Conservative 169; Mismatches 150; Indels 493; Gaps 222;
QY 1 YP--VV-L--AD--TSSSEDA-LN---I---SDKEKVAE-NK-----EK-HENIH 34
DB 458 YPLGLSVSYKQDPEWMSVS-DARLNGIRIRIELVNDPSDITEQNPVLVYRWEEVSQALH 516
QY 35 --SAMETSQ-DFKE--KK-T--AVIKEKV--VS--KNPVID--N--N-TSNEA--KIK 75
DB 517 QPKA-EP-QTELEADQELNLFSLREBPVQSIGLE-PD-DSNGHNDTDLJETDQIP 572
QY 76 EENENKS--QG--D-Y-----TDSF-----V--N-----KN--TE--N--P-KK 100
DB 573 EEEVEETIPEIPTVDFPPEDLTD-FYPKTARDKVEETNIVAIRLVKQLEVHRNASPSEQ 631

QY 101 E--DKVY-Y--IA-E-EKD-----KESGE-KA-I--KELSSLKNTK--VLVT----- 134
DB 632 ELAKYVGWGLANIEFPDYNPKEFSKERELKSLVTDKXYSMDKQSLTAYTDPSLIRQ 691
QY 135 -YD---R-INGSAIETTPD-----N-----LDK-IKQ-IE--GI--SSVEEA-QK-VOPMM 173
DB 692 MMDLEEDGFTGGKI-LDPSMGTCNFPAPMPKHLREKSELYGVELDTITGAIAKHLAP-- 748
QY 174 N-HARKKI--EVEEAT--D--Y-LKSI--NAPF-----GKU-----F----- 200
DB 749 NSHI--EIKGF-ETVANDNSFDL-VISNVFANIRIADRRYDRPYINHDFVAKSIDLL 804
QY 201 -DEGRNV-ISNI-DTGT-D-----Y--RH--K--AMRIDDA-KA-----S-----M-R 232
DB 805 HDG-GQVAL--ISSTGMDKRTENIIGDIRTTEFLGCVNLPDPAFAIAGTSVTTMLF 861
QY 233 F-K--K--EDL--KGT--DKN--YWLSDKI PHAF-----N-----Y--YNGGKI 263
DB 862 FOKHLDKGYVADDLAFSGSIRYDKSRIML--NPF--FDGHSYQVLGTVEVNFNGTIL 917
QY 264 TVEKYDDGRDYPD--PHGM-HI-AG--I-----LAGN--D--TBO--D--I-----KN----- 297
DB 918 SVKGTSD--DLIASVETALNHVKAPREIDRNEVITINPDVLTKOVNDTSIPAEMKRENLQY 975
QY 298 -FN--G--I--D--GI--APNAQIFSYKMYSD-AGSGF-AGDETWFH--AIED--S 336
DB 976 SFGQSGTYYRRDMKGRVGTKEEII-SY--YDDEG-NFKAMD-TR-HGQKQ--DRFNA 1028
QY 337 IK--HN--VDV--VS--G-FTGTGLVEEK--YMOA-IRALRK-AGIP-MV-VATGN- 380
DB 1029 LEVTDNLTALVDVYDDAKRQGF--KGYV-KKTVFYFAPL-SYKAVARIKGMVIDR--NA 1082
QY 381 YA-TSASSSMDLY--ANNHL--KM--T-DT-----G--N--YTR-----T-AAH 412
DB 1083 YOEVIATQRYDYIDKETFNHLGKLNRTYDSFVKHGYLNSAVNRNLPDSDDKYSLASL 1142
QY 413 ED-AI-----AV--AS-A-KNOTV-----B-----FDKVN--I--G-SESFXY-RNIG 444
DB 1143 EDESLDPSGKSVIYTTSLAEKLVREKEVKVYHTALDLNLSLADRGVDFAVMSI- 1201
QY 445 AF-PDKSKITTED-G--TKAPSK-L-KFYVIGQDQDLI-GLDLRKXIAMDRIYT 495
DB 1202 -YGV-ESQMTLIBELGDLIMPDEKYLNGELTYV--SR-ODPLSG-DVYTKLEVVD-LFV 1254
QY 496 K-DLKNA-EKKANDKG-----ARAIMVNTVNY--NNDNTELP-----A-----M 533
DB 1255 KQD--NODFMSHVAGLLEAIKPAR-ITLAD-IDYRIGSR--W--IPLAVYGRFAOETFM 1306
QY 534 G--YE-AD-EGTSQVPSISGDDGV-----KL-WNMIN-PDKTEVK--RNKED----- 575
DB 1307 GKAYELSDQE-VAT-VLEVPIDGVITYGSKPATYTSNATDRSLGVAASR---YDSRKI 1351
QY 576 FKQKL--EQYYP-IDMESF--NSNKNPYGD-E-----KE--I-D-FK-P-A--P----- 610
DB 1362 FENLINSNQ--PITTKQVNVGDKKQNTVDVETKTVLAKETHIQELFQGVAKPYEQOM 1419
QY 611 --PTDKELYKEDIIVPAGSTSW-G-PR-IDLLKPDVS-AP-GKN-----I---K-STLN- 654
DB 1420 IETPTNLYNR--TV--SKSYDSHLTID-GLAONISLRPHQNAIQRIVEERALLAH 1473
QY 655 -VINGKS-T-YG 663
DB 1474 EVGSGKTLTMLG 1485
RESULT 14
ABP28340
ID ABP28340 etandard; Protein; 2278 AA.
XX
AC ABP28340;
XX
DT 02-JUL-2002 (first entry)

XX Streptococcus polypeptide SEQ ID NO 5856.
XX
XX Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KM Streptococcus; Streptococcus pyogenes; antibacterial;
KM antiinflammatory; infection; vaccine; meningitis; gene therapy.
XX
OS Streptococcus agalactiae.
XX
XX WO200234771-A2.
XX
XX 02-MAY-2002.
XX
XX 29-OCT-2001; 2001WO-GB04789.
XX
XX 27-OCT-2000; 2000GB-0026333.
XX
XX 24-NOV-2000; 2000GB-0028727.
XX
XX 07-MAR-2001; 2001GB-0005640.
XX
XX (CHIR-) CHIRON SPA.
XX
XX (GENO-) INST GENOMIC RES.
XX
XX Telford J, Maignani V, Margalit Ros YI, Grandi G, Fraser C;
PI Tectel H;
XX
XX WPI; 2002-352536/38.
XX
XX N-PSDB; ABN68971.
XX
XX New Streptococcus protein for the treatment or prevention of infection
PT or disease caused by Streptococcus bacteria, such as meningitis, and
PT for detecting a compound that binds to the protein -
XX
XX
XX Claim 1; Page 3748; 4525dp; English.
XX
XX The invention relates to a protein (ABP25413-ABP30895) from group B
CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
CC the specification. The proteins have antibacterial and antiinflammatory
CC activity. (I), nucleic acids encoding (I), ABN68044-ABN71526 and
CC antibodies that bind (I) are used in the manufacture of medicaments for
CC the treatment or prevention of infection or disease caused by
CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
CC Nucleic acids encoding (I) are used to detect Streptococcus in a
CC biological sample. (I) is used to determine whether a compound binds to
CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
CC used as a vaccine or diagnostic composition. The disease caused by
CC Streptococcus that is prevented or treated may be meningitis. Nucleic
CC acid encoding (I) may be used to recombinantly produce (I) and may be
CC used in gene therapy. Antibodies to (I) are used for affinity
CC chromatography, immunoassays, and distinguishing/identifying
CC Streptococcus proteins.
XX
XX
XX Sequence 2278 AA;
SQ
Query Match 34.2%; Score 1506; DB 23; Length 2278;
Best Local Similarity 25.6%; Pred. No. 1.5e-15;
Matches 280; Conservative 169; Mismatches 150; Indels 493; Gaps 222;
QY 1 YP---VV-L-AD---TSSSEDA-IN---I---SDKEKVAE-NK-----EK-HENIH 34
DB 462 YPLGSLVYSGQDFEVMSSVS-DARLNGLRIRLAVDFSDIIRQNVPILVYRTVEEVSQALH 520
QY 35 --SMEIETSO-DFKE--KK-T-AVIKEKEV--VS--KNPVID--N--N-TSNEEA--KIX 75
DB 521 QPKA-EP-QTELEBADQELNLFSLFEEBPVOSIGLLE-PD-DSENGHNDTDLLEETDNOIP 576
QY 76 EENSNSK--QG---D-Y-----TDSF-----V--N-----KN--TE--N--P--KK 100
DB 577 EEEVETIPEIPEIYVDYFPEPDLTD-FYPKTARDKVEITNIVATRLVKNLEVEHRNAPSQ 635
QY 101 E-DKVV-Y--IA-E-FKD-----KESGE-KA-I--KELSLKNTK--VLYT----- 134
DB 636 ELAKYVGMGLANFPDYNPKFSKEREBELKSLVTDKEYSDMKOSSILTAAYTDSLIRQ 695

QY 135 -YD---R-IFNGSAIETTPD-----N-----LDK-IKQ-IE--GI--SSVERA-QK-VOPWM 173
DB 686 MMDKLERGFTGCKI-LDPSMGTFNFPAMPBPHLREKSELVGVELDTITGALAKLHP-- 752
QY 174 N-HARKEI-GVEBAI--D-Y-LKSI-NAPF-----GKN-----F----- 200
DB 753 NSHI--EIKGF-ETVAFNDNSFDL-VISNVFANIRIADRNPYMTIDYFVKXSLDLL 808
QY 201 -DGRGMV-ISNI-DTGT-D-----Y--RH--K--AMRIDDDA-KA-----S-----M-R 232
DB 809 HDG-GQVAI--ISSGTMDKRTENILQIRRTETFLGGVLPDSAFKALAGTSVTTDMIF 865
QY 233 F-K--K-----EDL--KGT--DKN--YMLSDKIPHAFF--N-----Y--NGSKI 263
DB 866 FOKHLDKGVAVDDLAFLSGSIRYDKDSRIML--NRY-FDGEVNSQVLGYEVBRNFGTL 921
QY 264 TVEKYDGRDYFD--PHGM-HI-AG-I-----LAGN-D--TRQ--D--I-----KN----- 297
DB 922 SVKGTSD--DLIASVETALNHVKAPREIDRNEVIINPVLTKQVADTSIPAMRENTLQY 979
QY 298 -FN--G--I---D--GI---APNQIFSYKMSD-AGSGF-AGDETMFH---AIED---S 336
DB 980 SFGYQSTVYVRDNGKIRVGTKEBI-SY--YVDEEG-NFKAMD-TK-HSOKOI-DRENA 1032
QY 337 IK--HN--VDV-VS--VSS-G-FTGTLVGEK--YMOA-IRALRK-AGIP-MV-VATGN- 380
DB 1033 LEVTDNTALDYYVDDDAKRGF--KGYV-KKTVEYEAFL-SYKEVARIKGVADR--VA 1086
QY 381 YA-TSASSSMDLV--ANNHL--KM--T-DT-----G--N--VTR-----T--AAH 412
DB 1087 YQEVIAIQRYVDYDDETENHLGKLNRTYDSFVKHYGYLNSAVNNHLPDSDDKYSILASL 1146
QY 413 ED-AI-----AV---AS-A-KNOGV-----E-----FDKVN--I--G-GSGFKY-RNG 444
DB 1147 EDESIDPGCKSVIYTKSLAFKALVRPREVKKCHTALDALSSLAGDKGVDFAYMGI- 1205
QY 445 AF-FDKSKITTNE-D--G--TRAPSK-L--KFVYIGKGDQDPLI-GLDLRGKIAVMDRIYT 495
DB 1206 -YQV-ESQMTLIEELGDLIMDPBKXNGELTYV--SR-QDPLSG-DVYTKLEVVVD-LFV 1258
QY 496 K-DLNA-FKAMDKG-----APALMVNTVNY--NRDWTLP-----A---M 533
DB 1259 KOD--NODPFMNSHYGLLEAIKPAR-ITLAD-IDRIGSR--W--IPLAVGKFAQETFM 1310
QY 534 G--YE-AD-EGTKSVFSGDDGV-----KL--NMNIN-PDKTEVK--RNNKE----- 575
DB 1311 GAYELSDQE-VAT-VLEVSPIDGVITYQSKFAITYSNATDRSLGVPSR---YDSGRKI 1365
QY 576 FKDKL--EQYVP-IDMESF-NSNKEPVGD-E-----KE--I-D-FK-F-A--P----- 610
DB 1366 FENLANSNQ--PTIKQYVEGDKKKNVDEKTTVLRABKETHLQGLFGQFVAKYEVQOM 1423
QY 611 --DTDKELYKEDIIVPAGTSM-G-P-R-IDLLKPDVS-AP-GKN---I---K-STLN- 654
DB 1424 IEDTYNRLYNR--TV---SKYDGSHTLTD-GLAONISLRPHQKNAIGRIVEKQALAH 1477
QY 655 -VINGKS--T-YG 663
DB 1478 EVSGSKTILMLG 1489
RESULT 15
AAB10667
ID AAB10667 standard; Protein; 2057 AA.
AAB10667;
AC AAB10667;
XX
XX
XX 19-JAN-2001 (first entry)
XX
XX L. mesenteroides alternan sucrose protein.
XX
XX Alternan sucrose; glucosyltransferase; fructose; cosmetic; foodstuff;

KW 8Yrup.
 XX Leuconostoc mesenteroides.
 OS DB19905069-A1.
 XX
 XX
 XX 10-AUG-2000.
 PD
 XX 08-FEB-1999; 99DE-1005069.
 XX
 XX 08-FEB-1999; 99DE-1005069.
 PR 08-FEB-1999; 99DE-1005069.
 XX
 XX (PLAN-) PLANTTEC BIOTECHNOLOGIE GMBH.
 PA (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
 PI Koesmann J, Weissh T, Quanz M, Knuth K,
 XX WPI: 2000-550294/51.
 DR N-PSDB; AAA97904.
 XX
 XX New nucleic acid encoding recombinant Leuconostoc mesenteroides
 PT alternan sucrose protein and methods of alternan and fructose
 PT production -
 XX
 PS Claim 1a; Page 30-36; 64pp; German.
 XX
 CC This invention describes a novel nucleic acid molecule (I) encoding an
 CC alternan sucrose (E.C. 2.4.1.140 - an enzyme, that belongs to the
 CC glucosyltransferase group) The recombinant, purified alternan sucrose
 CC gene is useful for the fermentative production of alternan (a
 CC carbohydrate) and/or fructose by secreting the enzyme into a
 CC saccharose-containing culture medium. Alternatively, the enzyme is
 CC contacted with a saccharose-containing solution. The alternan and/or
 CC fructose is then isolated from the medium. Cosmetic products or
 CC foodstuffs containing alternan can be produced. Recombinant production of
 CC alternan sucrose is advantageous as it provides a cost effective means of
 CC producing fructose for high fructose containing syrups, production of
 CC which previously has been achieved by costly production from maize
 CC starch. This sequence represents the Leuconostoc mesenteroides alternan
 CC sucrose protein which is described in the method of the invention.
 CC
 XX
 SO Sequence 2057 AA;
 Query Match 34.2%; Score 1502; DB 21; Length 2057;
 Best Local Similarity 24.3%; Pred. No. 1.1e-15;
 Matches 296; Conservative 150; Mismatches 155; Indels 617; Gaps 240;
 QY 1 YPV-VLAQT-SSSE---DAL-N-ISDKEK--V---AE---NKEKHNHSAM-E 38
 DB 63 VVPSTNDSILKQGTDFWYDSGNRV-D-QKTQIILTLAEQLKKNNEK--NL-SVISDD 117
 QY 39 TS-ODFKEKTAIVIKKEKVSKN-PVIDN---NTSBEAKI--KE-ENSKSGCD--YT 87
 DB 118 TSKKD-DE---NISKQTKIA--NOQTVDYAKGLTTSLSDPITGAYENHN--GYRVI 168
 QY 88 D-S-F-VN--KNTN-N-PKEDKVVY-I-ABFKDSEGEKAIKELSL--K---NTKVL- 132
 DB 169 DASGKQVTGQNTIDGNLQYFDN-GYVKGSPFRV-NG-KHIV-FDVTGKASNVIVN 224
 QY 133 -----Y-----TY--D---RI--F--NG-SAI--ETTPD-NL--DK--IKQIE-G 159
 DB 225 GKAGYDAQGNOLKKSIVYADSSGQTYFDGNGPLIGLQTI-DGNLQYFNGQGV-QIKGG 282
 QY 160 ISSV--BRAQKVOPMM-NHARK--EI--G--VE--EAI-DVLKSNAPF-----GKN---F 200
 DB 283 FQDVNNKRI-YFAPNTGN-AVANTETLNGKLQGHANDANGNVR--NA-FSDVAG-NTPYF 336
 QY 201 DGRG-WV--ISNIDTG-TDY---R-H--K--A-----NRIDD--A-YASM--RF--- 233
 DB 337 DANGVMLTGLQTI-SGKTYLVLEQGHLRKQYAGTFNNQFVFPADDTGAGKTAIEYQPDG 395
 QY 234 -----KKE-----D---LK---G-----TD--KN--YVL-S---DK---I- 251

DB 396 LVQSONENTPHNAKSYDKSSFENVQGYLTADTYRPTDILKNGDTWTASTETDMRELLM 455
 QY 252 ---P---HAFTYNN-----G-G-----K-----I-T-VE-----K---Y 268
 DB 456 TWMPDKQTOA-NYLNPFSSKGLGITTTTYTAATSOKTINDAFAVYQTAIBQOISLKSTEW 514
 QY 269 -D-D-----G-R-----D---Y---F-----DPH-----GMH--IAG--- 284
 DB 515 LRDAIDSFVKTQANWNNQTEDEAFDGLQMLQGGFLAVQODSSHPTPTDGSNNRL-GROP 573
 QY 285 I-LAG-NDTEODIKN--F---NGIDGIAFPAQT--FSYKMY-SDAGSGFAG--DETFFHA 332
 DB 574 INTDSKDT-TDKGSEFLLANDIDNSNPYQAEQMLMTHLYLNKFGS-IYGNNDNANFDG 631
 QY 333 IE-DSIKENV--DVVSVSSG--F-T--GT-----G-L-VGBKVV-----Q-AIALR 368
 DB 632 IRVDAVD-NVDADILKI-AGDYFKALYGTGKSDANANKHLSLID--NGKDPQTVNO--Q 686
 QY 369 KAG---IPM--VYAT--GN-YAT-SAS-S-S-W---D-LVAN---NHLKPTD-----TG 404
 DB 687 --GNAQLTMDYVTVTSQFNSL--THGANNRSMNMYPLDTGYLLNGDLN-KKIVDKNRNSG 742
 QY 405 N-VTR-----T-----A-AHE-DA---I--A-V--ASAKR-Q-TVEPKVNIGES 437
 DB 743 TLVNRIRANSQDTVIVPNYSFVRADYDAQPIPKAMIDHGIIRKMDPTFFDOLAQGM- 801
 QY 438 FKY--R-N-IGAF--F-D---KS---KITTED-----GT--KAP- 462
 DB 802 FYKQGNBPG-FKKTNDVNLPSAYVAMLTKQTVPRVYGYDMYLBEGQYMEKGTIYNPV 860
 QY 463 -SK-LK-FV-YI-GKQD---D---ODLI-G-LDL---R-GKIAVW--DRIYT---KD 497
 DB 861 ISALLKARIYVVG-QGTMAIDSGDKDKGERTDLTSVAFGR-GIMTSQTTQDMSQD 918
 QY 498 LKNA-----FKKAMDK-----G-A-----RAIMV-N-T-VNYN----- 523
 DB 919 YKQGGIVYGNPNPDKLNNDKITTLHMGVAKHQULYRA-LVISNDSGIDVDSDBKAPT 977
 QY 524 -R--DNWTEL---P---AW--G---YEADGKTS-QVPSIGDGSVKLMNNINP----- 562
 DB 978 LRTNDNG-DLIFPKTNTFVAQDGTITNYEM-KG--SLNAL-ISGYLGV--W--V-PVGAS 1027
 QY 563 ---DKK---TEVRANKED---F--KDKLE-QY-YPIDMESFNSN--KP-----NV- 598
 DB 1028 DQDARTVALESSSN--DSGVSHSNAALDSNVIY---EGF-SNQAAMTSPSGSTNV 1080
 QY 599 -GDE---KE-I-DRKAP-----DT-----DKEL--YK--E--DITV-PA-GS 627
 DB 1081 IATKAMLFKELGITSEFELAPQYRSSGDTYVCGMSFLDSFLANGYAFTRDGLFNNRADGN 1140
 QY 628 ---TSMGP---R--IDLILK-----PP-VGA-PCK-----N-IKST- 652
 DB 1141 PNPTRKGTODDLNNALEAHKNGMOAIDWPQOIALPKEKVVYATRVDERGNOLKDTD 1200
 QY 653 -LN---VINGKSTYV--Y 664
 DB 1201 FVNLVYVANTKSS-GVDY 1217
 Search completed: October 14, 2003, 16:30:18
 Job time : 98 secs

 pat

 US09590991

 dir

 runat_14102003_154349_14329

 printer

 0

```
</UL>
<LI><STRONG>html_save</STRONG>
<UL>
<LI>0
</UL>
<LI><STRONG>runat_14102003_154349_14329</STRONG>
<UL>
<LI>US-09-590-991-6
</UL>
<LI><STRONG>save</STRONG>
<UL>
<LI>Save As..
</UL>
</UL>
IN=/cgn2_1/USPTO_gpool/US09590991/runat_14102003_154349_14329/output, F=/cgn2_1/USPTO_gf
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: October 14, 2003, 16:26:41 ; Search time 48 Seconds
(without alignments)
2195.718 Million cell updates/sec

Title: US-09-590-991-6
Perfect score: 4398
Sequence: 1 YPVVLADTSSSEDALNISDR.....PGKNIKSTLVINGKSTVGY 664

Scoring table: BLOSUM30
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /SIDSL1/gcgdata/geneseq/geneeqp-emb1/AA1980.DAT.*
2: /SIDSL1/gcgdata/geneeq/geneeqp-emb1/AA1981.DAT.*
3: /SIDSL1/gcgdata/geneeq/geneeqp-emb1/AA1982.DAT.*
4: /SIDSL1/gcgdata/geneeq/geneeqp-emb1/AA1983.DAT.*
5: /SIDSL1/gcgdata/geneeq/geneeqp-emb1/AA1984.DAT.*
6: /SIDSL1/gcgdata/geneeq/geneeqp-emb1/AA1985.DAT.*
7: /SIDSL1/gcgdata/geneeq/geneeqp-emb1/AA1986.DAT.*
8: /SIDSL1/gcgdata/geneeq/geneeqp-emb1/AA1987.DAT.*
9: /SIDSL1/gcgdata/geneeq/geneeqp-emb1/AA1988.DAT.*
10: /SIDSL1/gcgdata/geneeq/geneeqp-emb1/AA1989.DAT.*
11: /SIDSL1/gcgdata/geneeq/geneeqp-emb1/AA1990.DAT.*
12: /SIDSL1/gcgdata/geneeq/geneeqp-emb1/AA1991.DAT.*
13: /SIDSL1/gcgdata/geneeq/geneeqp-emb1/AA1992.DAT.*
14: /SIDSL1/gcgdata/geneeq/geneeqp-emb1/AA1993.DAT.*
15: /SIDSL1/gcgdata/geneeq/geneeqp-emb1/AA1994.DAT.*
16: /SIDSL1/gcgdata/geneeq/geneeqp-emb1/AA1995.DAT.*
17: /SIDSL1/gcgdata/geneeq/geneeqp-emb1/AA1996.DAT.*
18: /SIDSL1/gcgdata/geneeq/geneeqp-emb1/AA1997.DAT.*
19: /SIDSL1/gcgdata/geneeq/geneeqp-emb1/AA1998.DAT.*
20: /SIDSL1/gcgdata/geneeq/geneeqp-emb1/AA1999.DAT.*
21: /SIDSL1/gcgdata/geneeq/geneeqp-emb1/AA2000.DAT.*
22: /SIDSL1/gcgdata/geneeq/geneeqp-emb1/AA2001.DAT.*
23: /SIDSL1/gcgdata/geneeq/geneeqp-emb1/AA2002.DAT.*
24: /SIDSL1/gcgdata/geneeq/geneeqp-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4398	100.0	664	22	AA848342
2	4398	100.0	2120	21	AA81710
3	4379	99.6	2140	24	ABU01020
4	4171	94.8	1007	19	AA61246
5	4171	94.8	1007	23	ABP54664
6	974	22.1	192	13	AA186159
7	899.5	20.5	1962	12	AA10559
8	889.5	20.2	1962	12	AA10557
9	889.5	20.2	1968	12	AA10941

10	888.5	20.2	1962	12	AA10558	Mutant protease (A
11	887.5	20.2	1974	12	AA10940	Mutant protease (d
12	885.5	20.1	1962	12	AA10560	Mutant protease (K
13	885.5	20.1	1962	12	AA10561	Mutant protease (N
14	885.5	20.1	1962	12	AA10563	Mutant protease (K
15	884	20.1	1959	12	AA10562	Mutant protease (d
16	866.5	19.7	1560	10	AA94145	S. cremoris protei
17	817.5	18.6	1946	19	AA47273	Lactobacillus bulg
18	804.5	18.3	1570	23	ABP296894	Streptococcus poly
19	804.5	18.3	1570	23	ABP30531	Streptococcus poly
20	804.5	18.3	1590	23	ABP28560	Streptococcus poly
21	802.5	18.2	1550	23	ABP30895	Streptococcus poly
22	793	18.0	1579	24	ABP71293	Group B Streptococ
23	786	17.9	885	23	ABP27332	Streptococcus poly
24	786	17.9	885	23	ABP29796	Streptococcus poly
25	773	17.6	1647	23	ABP25822	Streptococcus poly
26	768	17.5	1239	23	ABP25822	Streptococcus poly
27	767	17.4	1233	23	ABP25675	Streptococcus poly
28	767	17.4	1233	23	ABP30203	Streptococcus poly
29	707	16.1	806	13	AA27481	RP-III residual pr
30	701	15.9	690	11	AA804904	Residual protease-
31	691	15.7	690	18	AA16337	Bacillus subtilis
32	691	15.7	690	18	AA101787	Residual protease
33	686	15.6	815	23	ABP91806	Herbically activ
34	669.5	15.2	1167	18	AA22470	SCP12 peptidase (
35	668.5	15.2	1167	18	AA22469	SCP12 peptidase (
36	667.5	15.2	1164	21	AA22469	SCP12 peptidase (
37	665.5	15.1	1164	21	AA801263	SCP49 peptidase (
38	661.5	15.0	1150	21	AA801265	SCP49 peptidase (w
39	660.5	15.0	1150	21	AA801266	SCP49 peptidase (w
40	657.5	14.9	1181	21	AA801266	SCP49 peptidase (w
41	657.5	14.9	1181	21	ABP25823	Streptococcus poly
42	652.5	14.8	1134	23	ABP29883	Streptococcus poly
43	645.5	14.7	1090	23	ABP298458	Streptococcus poly
44	641	14.6	903	17	AA87007	Hyperthermostable
45	641	14.6	922	19	AA87372	Trimmed enzyme pro

ALIGNMENTS

RESULT 1	AA848342	standard; Protein; 664 AA.
AC	AA848342:	
XX		
DT	20-APR-2001	(first entry)
XX		
DE	S. pneumoniae Sp128 polypeptide.	
XX		
KW	Immunogenic; Sp128; Sp130; pneumococcal; otitis media; nasopharyngeal;	
KW	bronchial; lung; blood; infection; immune response; immunotherapy;	
KW	antibacterial; auditory; vaccine.	
XX		
OS	Streptococcus pneumoniae.	
XX		
PN	W0200076540-A2.	
XX		
PD	21-DEC-2000.	
XX		
PF	09-JUN-2000; 2000MO-US15925.	
XX		
PR	10-JUN-1999; 99US-0138453.	
XX		
PA	(MEDI-) MED IMMUNE INC.	
XX		
PI	Adamou JE, Choi GH;	
XX		
DR	WPI, 2001-112197/12.	
DR	N-PSDB; AAC84741.	
XX		
PT	New vaccine comprising Sp128 or Sp130 polypeptides, for treating and	

PT preventing pneumococcal infections, particularly infections caused by
 PT Streptococcus, e.g. otitis media, nasopharyngeal, bronchial, lung or
 PT blood infections

PS Claim 4; Page 47-50; 54pp; English.

XX The invention relates to novel immunogenic polypeptides, Sp128 and Sp130
 CC from S. pneumoniae. Vaccines comprising the polypeptides are useful for
 CC the treatment and prevention of pneumococcal infections, particularly
 CC infections caused by Streptococcus, such as otitis media, nasopharyngeal,
 CC bronchial, lung or blood infections. The antigens are used as immunogenic
 CC agents to stimulate an immune response. The antisera and antibodies may
 CC also be used in diagnosing and treating pneumococcal infections.
 CC Recombinant polypeptides serve as a mechanism for stimulating production
 CC of antibodies for use in passive immunotherapy, diagnostic reagents, and
 CC as reagents in other processes such as affinity chromatography. The
 CC present sequence represents the S. pneumoniae Sp128 polypeptide.

XX Sequence 664 AA;

Query Match 100.0%; Score 4398; DB 22; Length 664;
 Best Local Similarity 100.0%; Pred. No. 3.7e-60;
 Matches 664; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YPVVLADTSSSEDALNLSDEKVAENKEKHENHSAMETSODFKEKTAIVIKEKEVSKN 60
 DB 1 YPVVLADTSSSEDALNLSDEKVAENKEKHENHSAMETSODFKEKTAIVIKEKEVSKN 60
 OY 61 PVIDNNTSNEBAKIKEENSNKSGDYTDSFVNKNTENPKKEDKVYVIAEFKDSGEKAI 120
 DB 61 PVIDNNTSNEBAKIKEENSNKSGDYTDSFVNKNTENPKKEDKVYVIAEFKDSGEKAI 120
 OY 121 KELSLKNTKVLTYTDRIFNLSAIEETPDNLDKIKQIEGSSVERAOKVQPMNHARKKEI 180
 DB 121 KELSLKNTKVLTYTDRIFNLSAIEETPDNLDKIKQIEGSSVERAOKVQPMNHARKKEI 180
 OY 181 GVEBAIDVLKLSINAFPGKNFDGKGVNINIDTGTDRYRKHAMRIDDDAKASRFKEDLKG 240
 DB 181 GVEBAIDVLKLSINAFPGKNFDGKGVNINIDTGTDRYRKHAMRIDDDAKASRFKEDLKG 240
 OY 241 TDKNYLSDKIPHANFYNGGKITVEKYDDGRDYDPHGMHAGILAGNDTEODIKFNNG 300
 DB 241 TDKNYLSDKIPHANFYNGGKITVEKYDDGRDYDPHGMHAGILAGNDTEODIKFNNG 300
 OY 241 TDKNYLSDKIPHANFYNGGKITVEKYDDGRDYDPHGMHAGILAGNDTEODIKFNNG 300
 DB 241 TDKNYLSDKIPHANFYNGGKITVEKYDDGRDYDPHGMHAGILAGNDTEODIKFNNG 300
 OY 301 IDGIANPNOISYKMYSDAGSGPAGDETFPHAIEDSIRKNDVYVSSGFTGTGLVGEKY 360
 DB 301 IDGIANPNOISYKMYSDAGSGPAGDETFPHAIEDSIRKNDVYVSSGFTGTGLVGEKY 360
 OY 361 WQAIKALKRAGIPMVVATGNTVATSSASSSSWDLVANNHLMKMTDTGNVTRTAHEDAIIVAS 420
 DB 361 WQAIKALKRAGIPMVVATGNTVATSSASSSSWDLVANNHLMKMTDTGNVTRTAHEDAIIVAS 420
 OY 421 AKNQVEEDKNIIGESFRYRNIGAFPDKSKITTNEDGTAKPSKLFYIGGQDDILIG 480
 DB 421 AKNQVEEDKNIIGESFRYRNIGAFPDKSKITTNEDGTAKPSKLFYIGGQDDILIG 480
 OY 481 LDLRGKIIVMBRIYKDKLNAFKKAMDGARAIVVNVVYNNRNMTLPMAGEADEG 540
 DB 481 LDLRGKIIVMBRIYKDKLNAFKKAMDGARAIVVNVVYNNRNMTLPMAGEADEG 540
 OY 541 TKSQVFSISGDGVLAMNINPDKTEVYRNKKEDFKOLEQYPIIDMESFNSKNPNVGD 600
 DB 541 TKSQVFSISGDGVLAMNINPDKTEVYRNKKEDFKOLEQYPIIDMESFNSKNPNVGD 600
 OY 601 EKEIDFKAPDTEKELYEDIIIVPAGSTSWGPRIDLLKPDVSAFGKNIKSTLVNINGS 660
 DB 601 EKEIDFKAPDTEKELYEDIIIVPAGSTSWGPRIDLLKPDVSAFGKNIKSTLVNINGS 660
 OY 661 TYGY 664
 DB 661 TYGY 664

RESULT 2
 ID AAY81710 standard; Protein: 2120 AA.

XX AAY81710;

DT 02-JUN-2000 (first entry)

XX Streptococcus pneumoniae protein sequence ID3.

XX Streptococcus pneumoniae infection; immunogen; antigen; diagnosis; AIDS;
 KW bacterial pneumonia; asplenia; heart disease; lung disease; alcoholism;
 KW kidney disease; diabetes; immunosuppressive disorder; otitis media;
 KW pneumococcal septicemia; sinusitis; meningitis; therapy.

XX Streptococcus pneumoniae.

XX WO200006738-A2.

XX 10-FEB-2000.

XX 27-JUL-1999; 99WO-GB02452.

XX 27-JUL-1998; 98GB-0016336.

XX 19-MAR-1999; 99US-0125329.

XX (MICR-) MICROBIAL TECHNICS LTD.

XX Le Page RWF, Wells JM, Hamifly SB, Hansbro PM;

XX WPI; 2000-1995301/17.

XX N-PSDB; AAZ91806.

XX Streptococcal proteins and polynucleotides useful for diagnosis,

XX treatment and prophylaxis of bacterial infections

XX Claim 2; Page 41-42; 76pp; English.

XX This sequence represents a Streptococcus pneumoniae protein of the
 CC invention. The proteins (or their homologues, derivatives and/or
 CC fragments) are useful as immunogens or antigens. Immunogenic or antigenic
 CC compositions comprising the proteins are useful as vaccines and also in
 CC diagnostic assays. The sequences are useful for the detection or
 CC diagnosis of S. pneumoniae infection, by contacting a sample to be tested
 CC with them. Agents capable of antagonising, inhibiting or interfering with
 CC the function or expression of the protein or polypeptide are useful in
 CC medical compositions in the treatment or prophylaxis of S. pneumoniae
 CC infection. As the sequences can be used to treat S. pneumoniae infection,
 CC they can be used to treat bacterial pneumonia, which has high rates in
 CC young children, the elderly, and in patients with predisposing conditions
 CC such as asplenia, heart, lung and kidney disease, diabetes, alcoholism,
 CC or with immunosuppressive disorders, especially AIDS. They can also be
 CC used to treat pneumococcal septicemia, otitis media, sinusitis, and
 CC meningitis.

XX Sequence 2120 AA;

Query Match 100.0%; Score 4398; DB 21; Length 2120;
 Best Local Similarity 100.0%; Pred. No. 6.4e-59;
 Matches 664; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YPVVLADTSSSEDALNLSDEKVAENKEKHENHSAMETSODFKEKTAIVIKEKEVSKN 60
 DB 1 YPVVLADTSSSEDALNLSDEKVAENKEKHENHSAMETSODFKEKTAIVIKEKEVSKN 60
 OY 61 PVIDNNTSNEBAKIKEENSNKSGDYTDSFVNKNTENPKKEDKVYVIAEFKDSGEKAI 120
 DB 61 PVIDNNTSNEBAKIKEENSNKSGDYTDSFVNKNTENPKKEDKVYVIAEFKDSGEKAI 120
 OY 121 KELSLKNTKVLTYTDRIFNLSAIEETPDNLDKIKQIEGSSVERAOKVQPMNHARKKEI 180
 DB 121 KELSLKNTKVLTYTDRIFNLSAIEETPDNLDKIKQIEGSSVERAOKVQPMNHARKKEI 180

QY 181 GVEAIDYLSKINAPFGKNGPGRGVISNIDTGDYRHKAMRIDDDAKASRFKKEDLKG 240
DB 181 GVEAIDYLSKINAPFGKNGPGRGVISNIDTGDYRHKAMRIDDDAKASRFKKEDLKG 240
QY 241 TDKNYMLSDKI PHAFNYNGSKITVEKYDDGRDYFDPHGMIAGI LAGNTEODIKNPFG 300
DB 241 TDKNYMLSDKI PHAFNYNGSKITVEKYDDGRDYFDPHGMIAGI LAGNTEODIKNPFG 300
QY 301 IDGIAPNAQIFSYKMYSDAGSGFAGDETFMFAIDSIKHNVDVSVSSGFTGLVGEKY 360
DB 301 IDGIAPNAQIFSYKMYSDAGSGFAGDETFMFAIDSIKHNVDVSVSSGFTGLVGEKY 360
QY 361 MQAIRALRKAGIPMVVATGNYATSSASSSMDLVANNHLKMTDTGNVTRTAHEDAI AVAS 420
DB 361 MQAIRALRKAGIPMVVATGNYATSSASSSMDLVANNHLKMTDTGNVTRTAHEDAI AVAS 420
QY 421 AKNOTVEFDKNIGESFKYRNIGAFDPSKITTNEDGTAPSLKKEVYIGKGDODLIG 480
DB 421 AKNOTVEFDKNIGESFKYRNIGAFDPSKITTNEDGTAPSLKKEVYIGKGDODLIG 480
QY 481 LDLRGKIAVMDRIYTKDLKNAFKKAMDGARAIVVNTVYNNRDNMTLPRMGYEADG 540
DB 481 LDLRGKIAVMDRIYTKDLKNAFKKAMDGARAIVVNTVYNNRDNMTLPRMGYEADG 540
QY 541 TKSGVPSISGDDGYKLMNMINPDKTEYKRNKEDFKDKLEQYYPIDMESFNSKPNVGD 600
DB 541 TKSGVPSISGDDGYKLMNMINPDKTEYKRNKEDFKDKLEQYYPIDMESFNSKPNVGD 600
QY 601 EKEIDFPAPPTDKELYKEDIIVPAGSTSWGPRIDLKLPVSAFGKIKSTLAVINGKS 660
DB 601 EKEIDFPAPPTDKELYKEDIIVPAGSTSWGPRIDLKLPVSAFGKIKSTLAVINGKS 660
QY 661 TYGY 664
DB 661 TYGY 664

RESULT 3
ABU01020
ID ABU01020 standard; Protein: 2140 AA.

XX ABU01020;

DT 11-FEB-2003 (first entry)

DE S. pneumoniae type 4 strain protein from coding region #590.

XX Bacterial meningitis; pneumonia; sepsis; otitis media;
KM ear infection; antiinflammatory; antibacterial; immunostimulant;
KM auditory; respiratory; gene therapy; vaccine.

XX Streptococcus pneumoniae type 4 strain.

XX MO200277021-A2.

XX 03-OCT-2002.

XX 27-MAR-2002; 2002MO-IB02163.

XX 27-MAR-2001; 2001GB-0007658.

XX (CHIR-) CHIRON SPA.
PA (GENO-) INST GENOMIC RES.

XX Maignant V, Tectelin H, Fraser C;

XX WPI; 2003-040579/03.

XX N-PSDB; ABX06302.

PT New proteins and nucleic acid molecules from Streptococcus pneumoniae,
PT useful as medicaments for treating or preventing a disease or infection
PT due to streptococcus bacteria, such as pneumonia, sepsis, otitis media
PT or ear infection

XX Claim 1; SEQ ID No 1180; 56pp; English.

PS The invention relates to a protein comprising or having at least 50%
XX identity to any of the 2469 amino acid sequences, identified in the
CC specification (available on a computer readable format), or its fragment,
CC expressed from 2469 of 2489 identified DNA coding regions from the
CC Streptococcus pneumoniae type 4 strain genomic sequence appearing as
CC A8556454. Also included are an antibody which binds one of the
CC proteins, treating a patient by administering the protein, DNA or
CC antibody (in a composition), a kit comprising first and second primers,
CC which are the nucleic acid cited above or fragments between nucleotides
CC 8-100 of a sequence not defined in the specification, for amplifying a
CC target sequence contained within a Streptococcus nucleic acid sequence,
CC where the first primer is substantially complementary to the target
CC sequence and the second primer is substantially complementary to the
CC complement of the target sequence, and where the parts of the primers
CC having substantial complementarity define the termini of the target
CC sequence to be amplified, assay comprising contacting a test compound
CC with the protein, and determining whether the test compound binds to the
CC protein and a Streptococcus pneumoniae bacterium, where one or more
CC genes encoding the proteins has been rendered inactive. The proteins,
CC nucleic acid molecules, antibody and compositions are useful as
CC medicaments for treating or preventing a disease or infection due to
CC streptococcus bacteria, particularly S. pneumoniae, such as pneumonia,
CC sepsis, otitis media or ear infection. They are also useful in developing
CC vaccines, diagnostics and antibiotics. The methods are useful for
CC identifying immunodominant proteins. The present sequence is one of
CC the 2469 proteins expressed by the identified coding regions from the
CC genomic sequence.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pcr_sequences.

XX Sequence 2140 AA;

Query Match 99.6%; Score 4379; DB 24; Length 2140;

Best Local Similarity 99.8%; Pred. No. 1,3e-58;

Matches 662; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 PVVLAIDTSSSEDALNIDSKKVAENKKEHNISAMETSGDPFKKTAIVIKERKVSKNP 61
DB 23 EVVLADTSSSEDALNIDSKKVAENKKEHNISAMETSGDPFKKTAIVIKERKVSKNP 62
QY 62 VVNNNTSNEBAKIKERKSNKSGDYTDSPFNKNTENPKKEDKVVYIAEPDKEGGEKAIK 121
DB 83 VVNNNTSNEBAKIKERKSNKSGDYTDSPFNKNTENPKKEDKVVYIAEPDKEGGEKAIK 142
QY 122 ELSSLKNTKVLKYRDIYFNGSAIETTPDNDKIKOIEGSSVERAOKVOPMMHARKKIG 181
DB 143 ELSSLKNTKVLKYRDIYFNGSAIETTPDNDKIKOIEGSSVERAOKVOPMMHARKKIG 202
QY 182 VBEAIDYLSKINAPFGKNGPGRGVISNIDTGDYRHKAMRIDDDAKASRFKKEDLKG 241
DB 203 VBEAIDYLSKINAPFGKNGPGRGVISNIDTGDYRHKAMRIDDDAKASRFKKEDLKG 262
QY 242 DKNYMLSDKI PHAFNYNGSKITVEKYDDGRDYFDPHGMIAGI LAGNTEODIKNPFG 301
DB 263 DKNYMLSDKI PHAFNYNGSKITVEKYDDGRDYFDPHGMIAGI LAGNTEODIKNPFG 322
QY 302 DGIAPNAQIFSYKMYSDAGSGFAGDETFMFAIDSIKHNVDVSVSSGFTGLVGEKY 361
DB 323 DGIAPNAQIFSYKMYSDAGSGFAGDETFMFAIDSIKHNVDVSVSSGFTGLVGEKY 382
QY 362 MQAIRALRKAGIPMVVATGNYATSSASSSMDLVANNHLKMTDTGNVTRTAHEDAI AVAS 421
DB 383 MQAIRALRKAGIPMVVATGNYATSSASSSMDLVANNHLKMTDTGNVTRTAHEDAI AVAS 442
QY 422 KNOTVEFDKNIGESFKYRNIGAFDPSKITTNEDGTAPSLKKEVYIGKGDODLIG 481
DB 443 KNOTVEFDKNIGESFKYRNIGAFDPSKITTNEDGTAPSLKKEVYIGKGDODLIG 502

```

Qy 482 DLRGKIAVMDBR1YTKDLKNAFKKAMDKGARAIMVNTVYNNRDNMTLPAWGTEADSGT 541
Db 503 DLRGKIAVMDBR1YTKDLKNAFKKAMDKGARAIMVNTVYNNRDNMTLPAWGTEADSGT 562
Qy 542 KSOVFSISGDDGVKLMNNINPDKTEVGRNKKEDFKDLKLEQYYPIDMESFNSKKNVDE 601
Db 563 KSOVFSISGDDGVKLMNNINPDKTEVGRNKKEDFKDLKLEQYYPIDMESFNSKKNVDE 622
Qy 602 KEIDFKFAPDPTDKELYKEDIIVPAGSTSMGPRIIDLKPDVSAFGKNIKSTLNVINGKST 661
Db 623 KEIDFKFAPDPTDKELYKEDIIVPAGSTSMGPRIIDLKPDVSAFGKNIKSTLNVINGKST 682
Qy 662 YGY 664
Db 683 YGY 685

```

RESULT 4
AAM61246
ID AAM61246 standard; Protein; 1007 AA.

XX AC AAM61246;
XX DT 02-OCT-1998 (first entry)
XX DE Streptococcus pneumoniae SPI22 protein.

XX KM Streptococcus pneumoniae; antigen; vaccine; infection; diagnosis;
XX KN detection; pneumonia; otitis media; meningitis.

XX OS Streptococcus pneumoniae.

XX PN MO9818930-A2.

XX PD 07-MAY-1998.

XX PF 30-OCT-1997; 97WO-US19422.

XX PR 31-OCT-1996; 96US-0029960.

XX PA (HUMA-) HUMAN GENOME SCT INC.

XX PI Choi GH, Hromockyj A, Johnson LS, Kunsch CA;

XX DR WPI, 1998-272224/24.

XX N-PSDB; AAV27431.

XX PT Nucleic acid encoding antigenic peptide(s) from Streptococcus
XX PT pneumoniae - or their epitope-containing fragments, useful in
XX PT protective or therapeutic vaccines, and for diagnosis

XX PS Claim 11; Page 92-93; 118pp; English.

XX CC The present sequence represents a protein from Streptococcus pneumoniae.
XX CC The nucleic acid sequence encoding the Streptococcus pneumoniae protein
XX CC can be useful in vaccines for inducing protective antibodies against
XX CC Streptococcus pneumoniae, for treatment or prevention of infection e.g.
XX CC pneumonia, otitis media or meningitis. Probes based on the nucleic acid
XX CC are used to detect Streptococcus infection (by usual hybridisation or
XX CC amplification methods), also for isolating Streptococcus genes or their
XX CC allelic variants. The protein can be used similarly to detect specific
XX CC antibodies in standard immunoassays, especially for diagnosing or
XX CC monitoring infections. Antibodies which bind the protein are used to
XX CC detect corresponding antigens, to purify the protein and for passive
XX CC immunisation (optionally coupled to a toxin). Vaccines are administered,
XX CC e.g. by injection, orally or through the skin, typically at 0.01-1000
XX CC (especially 10-300) mu g/ml per dose.

XX SQ Sequence 1007 AA;

Query Match 94.8%; Score 4171; DB 19; Length 1007;
Best Local Similarity 100.0%; Pred. No. 3.8e-56;
Matches 627; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 38 ETSODFEKKTAVIKEKEVVSKNPVIDNNTSNEBAKIKEENSNSKOGDYTDSFVNKNTEN 97
Db 1 ETSODFEKKTAVIKEKEVVSKNPVIDNNTSNEBAKIKEENSNSKOGDYTDSFVNKNTEN 60
Qy 98 PKKEDKVVYIAEFKOKSGEKAILELSLKNTKLYLYYDRIFNGSAIETTDDNLDKIKOI 157
Db 61 PKKEDKVVYIAEFKOKSGEKAILELSLKNTKLYLYYDRIFNGSAIETTDDNLDKIKOI 120
Qy 158 EGISSVRAOKVQPMNMHARKEIGVEEAIDVLKSIAPFGNPDGRGVNSIDTGDYR 217
Db 121 EGISSVRAOKVQPMNMHARKEIGVEEAIDVLKSIAPFGNPDGRGVNSIDTGDYR 180
Qy 218 HKAMRIDDDAKASRFKKEDLKGTDKNYWLSDKIPHAFNYNGSKITVEKYDDGRDYDP 277
Db 181 HKAMRIDDDAKASRFKKEDLKGTDKNYWLSDKIPHAFNYNGSKITVEKYDDGRDYDP 240
Qy 278 HGWHIAGILAGNDTEODIKNFNGIDGIAIPNAQITSYKYSIDAGSGFAGDETFHAIEDSI 337
Db 241 HGWHIAGILAGNDTEODIKNFNGIDGIAIPNAQITSYKYSIDAGSGFAGDETFHAIEDSI 300
Qy 338 KHNVDVSVSSGFTGTGLVGEKYNQAIKALRKAGIPMVVATGNATYATSSSSMDLVANNH 397
Db 301 KHNVDVSVSSGFTGTGLVGEKYNQAIKALRKAGIPMVVATGNATYATSSSSMDLVANNH 360
Qy 398 LKMTDTGNVTTAAHEDAIIVASAKNOTVEFDKYNIGESPKYBNIGAFPDKXITTNED 457
Db 361 LKMTDTGNVTTAAHEDAIIVASAKNOTVEFDKYNIGESPKYBNIGAFPDKXITTNED 420
Qy 458 GTKAPSKLKFFYIGKGDODLIGLDLKGKIAVMDBR1YTKDLKNAFKKAMDKGARAIMVNVN 517
Db 421 GTKAPSKLKFFYIGKGDODLIGLDLKGKIAVMDBR1YTKDLKNAFKKAMDKGARAIMVNVN 480
Qy 518 TVNYNNDNMTELPAWGTEADSGTGSQVFSISGDDGVKLMNNINPDKTEVGRNKKEDFK 577
Db 481 TVNYNNDNMTELPAWGTEADSGTGSQVFSISGDDGVKLMNNINPDKTEVGRNKKEDFK 540
Qy 578 DKLKQYYPIDMESFNSKKNVNGDEKEIDFKFAPDPTDKELYKEDIIVPAGSTSMGPRIIDL 637
Db 541 DKLKQYYPIDMESFNSKKNVNGDEKEIDFKFAPDPTDKELYKEDIIVPAGSTSMGPRIIDL 600
Qy 638 LKPDVSAFGKNIKSTLNVINGKSTYGY 664
Db 601 LKPDVSAFGKNIKSTLNVINGKSTYGY 627

```

RESULT 5
ABP54664
ID ABP54664 standard; Protein; 1007 AA.

XX AC ABP54664;

XX DT 04-SEP-2002 (first entry)

XX DE S. pneumoniae SPI22 protein sequence SEQ ID NO:216.

XX KM Streptococcus pneumoniae; epitope; vaccine; antigenic protein;
XX KN antibacterial; Streptococcal infection; detection.

XX OS Streptococcus pneumoniae.

XX PN US2002061545-A1.

XX PD 23-MAY-2002.

XX PF 22-JAN-2001; 2001US-0765272.

XX PR 30-OCT-1997; 97US-0961083.

XX PA (CHOI/) CHOI G H.
XX PA (KUNS/) KUNSCH C A.
XX PA (BARA/) BARASH S C.
XX PA (DILL/) DILLON P J.

PA (DOUG/) DOUGHERTY B.
 PA (FRAN/) FANNON M R.
 XX (ROSE/) ROSEN C A.
 PI Choi GH, Kunsch CA, Barash SC, Dillon PJ, Dougherty B, Fannon MR,
 PI Rosen CA;
 XX WPI; 2002-4/9261/51.
 DR N-PSDB; ABQ84899.
 XX
 PT New Streptococcus pneumoniae antigens, useful for detecting
 PT Streptococcus and for preventing or attenuating disease caused by
 PT Streptococcus infection -
 XX
 XX Claim 11; Page 50; 70pp; English.
 XX
 CC ABQ84792 to ABQ84904 represents nucleic acids which encode the
 CC Streptococcus pneumoniae antigens given in ABP54557 to ABP54669.
 CC The S. pneumoniae antigens have antibacterial activity and can be
 CC used in vaccines. The S. pneumoniae antigens can also be used to
 CC prevent or attenuate a Streptococcal infection in an animal. The
 CC polynucleotides encoding the S. pneumoniae antigens can be used to
 CC detect Streptococcus nucleic acids. ABQ84905 to ABQ85130 represent
 CC primers used in the cloning of S. pneumoniae ORFs (open reading frames)
 CC which are used in an example from the present invention.
 CC
 XX
 SQ Sequence 1007 AA;
 Query Match 94.8%; Score 4171; DB 23; Length 1007;
 Best Local Similarity 100.0%; Pred. No. 3.8e-56;
 Matches 627; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 38 ETSQDFEKKTAIVKEKEVSKNPVINDNTSNEBAKIKEENSKNSQGDYDTSFVNKNTEN 97
 DB 1 ETSQDFEKKTAIVKEKEVSKNPVINDNTSNEBAKIKEENSKNSQGDYDTSFVNKNTEN 60
 QY 98 PKKEDKVVYIAEFKDKESGEKAIKEISLKNKTKVLYYDRIENGSAIETTPDNLKIKOI 157
 DB 61 PKKEDKVVYIAEFKDKESGEKAIKEISLKNKTKVLYYDRIENGSAIETTPDNLKIKOI 120
 QY 158 EGISSVVERAKVQPMNHAKREIGVEBALDYLSINAPFGKPNFGKMTVSNIDTGTDR 217
 DB 121 EGISSVVERAKVQPMNHAKREIGVEBALDYLSINAPFGKPNFGKMTVSNIDTGTDR 180
 QY 218 HKAMRIDDDAKASMRFKKEDLKTGDKNYMLSDKIPHAFNYYNGGKIIVKXKDDGRDYEDP 277
 DB 181 HKAMRIDDDAKASMRFKKEDLKTGDKNYMLSDKIPHAFNYYNGGKIIVKXKDDGRDYEDP 240
 QY 278 HGHMIAIGILAGNDTEQDINKFNIGIDGIAAPNAQIFSYKMSDAGSGFAGDETFPHAIEDSI 337
 DB 241 HGHMIAIGILAGNDTEQDINKFNIGIDGIAAPNAQIFSYKMSDAGSGFAGDETFPHAIEDSI 300
 QY 338 KHAVDVVSVSSGFGTGLNGEKYQATRALRKAGIPNVVATGNVATGATSSSSMDLVANNH 397
 DB 301 KHAVDVVSVSSGFGTGLNGEKYQATRALRKAGIPNVVATGNVATGATSSSSMDLVANNH 360
 QY 398 LKMTDIGNVTRTAHEDAIASAANKQTFEFDKNIGESPKYKNICAFPDKSKITTNEED 457
 DB 361 LKMTDIGNVTRTAHEDAIASAANKQTFEFDKNIGESPKYKNICAFPDKSKITTNEED 420
 QY 458 GTKAPSLKLFVYIGKGGDQDLIGDLRGKIAVMDRITTKDLKNAFKKAMDKGARAIWVN 517
 DB 421 GTKAPSLKLFVYIGKGGDQDLIGDLRGKIAVMDRITTKDLKNAFKKAMDKGARAIWVN 480
 QY 518 TUNVYVNDNNTELPAWGYEADGEGTSQVFSISGDDGYKLMNMINPDCKTEYKRNKEDFK 577
 DB 481 TUNVYVNDNNTELPAWGYEADGEGTSQVFSISGDDGYKLMNMINPDCKTEYKRNKEDFK 540
 QY 578 DKLEQVYPIIMESFNKPNVGEKEIDFKFAPDTEKELYKEDIIVPAGSTSMGPRIDL 637
 DB 541 DKLEQVYPIIMESFNKPNVGEKEIDFKFAPDTEKELYKEDIIVPAGSTSMGPRIDL 600
 QY 638 LKPDVSAFGKNIKSTLVINGKSTYGY 664

DB 601 LKPDVSAFGKNIKSTLVINGKSTYGY 627.
 RESULT 6
 ID AAY86159 standard; Protein; 192 AA.
 AC AAY86159;
 DT 10-APR-2000 (first entry)
 DE S. pneumoniae derived protein #368.
 KM Treatment; prevention; disease; diagnosis; gene therapy; screening;
 KM bacterial; antimicrobial; antibiotic; pathogenesis; infection.
 OS Streptococcus pneumoniae.
 PN WC9806734-A1.
 PD 19-FEB-1998.
 PF 15-AUG-1997; 97WC-US14436.
 PR 16-AUG-1996; 96US-0024022.
 PA (SMIK) SMITHKLINE BEECHAM CORP.
 PI Black MT, Hodgson JE, Knowles DJC, Lonetto MA, Nicholas RO,
 PI Scodola RK;
 DR WPI; 1998-159452/14.
 DR N-PSDB; AA296473.
 PT Streptococcus pneumoniae proteins and related DNA - useful for
 PT screening compounds for antibacterial activity
 PS Claim 5; Page 617; 640pp; English.
 CC This invention describes novel isolated Streptococcus pneumoniae
 CC polynucleotides (see AA296173-296494) and their encoded proteins (see
 CC AA296173-296494). The DNA, vectors and host cells described in the
 CC method of the invention are useful for the recombinant expression of the
 CC polypeptides. The polypeptides are useful for treatment or prevention of
 CC disease, or diagnosis of disease related to expression or activity of
 CC such a polypeptide. They can also be used to screen for compounds which
 CC interact with and inhibit or activate such a polypeptide. The
 CC polypeptides (or DNA encoding them, via gene therapy) are also useful
 CC for inducing an immunological response in a mammal. The antagonists are
 CC useful to inhibit such bacterial polypeptides. The polypeptides are
 CC particularly useful to identify antimicrobial compounds and antibiotics.
 CC They are also useful to determine their role in pathogenesis of
 CC infection, dysfunction and disease.
 CC
 XX
 SQ Sequence 192 AA;
 Query Match 22.1%; Score 974; DB 19; Length 192;
 Best Local Similarity 94.7%; Pred. No. 1.2e-07;
 Matches 161; Conservative 2; Mismatches 7; Indels 0; Gaps 0;
 QY 2 PVVLAATSSSEDAINTDKKVAENKKEHINISAMETSODPKKTAIVKEKEVSKNP 61
 DB 23 EVLADTSSSEDAINTDKKVAENKKEHINISAMETSODPKKTAIVKEKEVSKNP 82
 QY 62 VIDNNTSNEBAKIKEENSKNSQGDYDTSFVNKNTENPKKEDKVVYIAEFKDKESGEKAIK 121
 DB 83 VIDNNTSNEBAKIKEENSKNSQGDYDTSFVNKNTENPKKEDKVVYIAEFKDKESGEKAIK 142
 QY 122 ELGSLKNTKVVLYYDRIENGSAIETTPDNLKIKOIEGISSVERAKVNP 171
 DB 143 QLSLKNKTKVLYYDRIENGSAIETTPDNLKIKOIEGISSVERAKVNP 192

PT Mutant protease gene(s) and protease(s) - derived from type I and
PT III protease genes from lactococcal strains, used in fermentation
PT foodstuffs and flavourings

PS Diaclosure; Fig 1(1-7)+5(b); 29 pp; English.

XX The mutant protease A137G/K138D having new cleavage specificities is
CC obtained by replacing two amino acids.
CC This mutant may then be used to prepare hybrid proteases,
CC the fusion being between a type I and a type III protease of
CC L. lactis M92 and SK11.
CC The product has modified properties, e.g. thermostability,
CC alkaline/acid pH stability, oxidative stability, autolysis etc.,
CC compared to the parent protease(s). The proteases can be used for
CC preparing products (butter cheese, human and animal foodstuffs)
CC prepared with the aid of lactic acid bacteria.
CC See also AAQ10411-17 and AAQ10870-71.

XX Sequence 1962 AA;

Query Match 20.2%; Score 889.5; DB 12; Length 1962;

Best Local Similarity 25.8%; Pred. No. 0.00081;
Matches 179; Conservative 132; Mismatches 248; Indels 135; Gaps 30;

QY 1 YPV--VLADTSSSEDALNISDEKVAENKEHENIHSAMETSODPFKEKTAIVIKEKVS 58
DB LPVGEIOAKAISOQTGSSILANTVTAA-----TAKQAATDTTAAT----- 65
QY KNPIVDNNTSNEAKIKEENSKNSQGDYTDSPFNKNTENPKEDKVVYIAFPKESGEK 118
DB 66 -NQAIAQLAAKIDYNKLNKVQODIYDVIVQM-SAAPASENGILRTDYSSTAETQOE 123
QY 119 AIKELSLAKTKVLY-----TYDRIFNQSAIETTPDNLDKIQIEGISSVERAOK 168
DB 124 TNKYIAQAQSKRAVEQVTOOTAGESTGYVNGSTKRVVDIKLQIACVKTITLAKV 183
QY 169 VQPMNHARKEIGVEALDYKLSINAPFGKNFDRGNVINSIDTGTYYRHAKMRIDDAK 228
DB 184 YPPTDAKANSMAVQAVMSNYK-----YKGEQTVSVISDGIPTHKDKMLSDDKD 234
QY 229 ASMRFKEDL-KGTDK--NMYLSDKI PHAENV- YNGKTIYKVDGRDYDPHGMHIA 283
DB 235 V--KLTSDVEKFTDYVHGKRYFNSKVPYGCNVADNDITIT---DDKVD--EQHGMHVA 286
QY 284 GILAGNTEODIKFNGIDGIAFPAQJFSYKMYSDAG--SGPAGDETWFHAIEDSIKXNV 342
DB 287 GIIAGNTEGDDPA--KSVGVAPRAQLAKKVPFNSDTSPTGSAIVYSAIEDSAKIGAD 344
QY 343 VSVVSSG-FTGTGLVGERKQWAIPLAKKAGIPMVAVATGNVATSASSSSWDLVANNHLM 401
DB 345 VLNNSLGNSNGNQTLEDEPBLAAVONANESGTAFAVISANGSTGSAATEG--VNDYDYLQ 402
QY 402 DTGNVTATAHEDLIAVASAKNOVEEDKNI-----GGESEFKY--RNIGAFPPDKSK 451
DB 403 DNEWVGSPGSRGATVYASAEENTVITQAVTITDGTGLQLOPEFTIQLSSHPFTGSPQCK 462
QY 452 ITTNEDETKAPSKLKFYIGKQDODLIGDLRGKIVAMDR- VTKLAKNAFKKAMDKGA 510
DB 463 FYIVKDSGN-----LSKGLADDTTA--DAKSKIAIVKRGKSPFDKOKIAQAA--GA 511
QY 511 RAINVNTVTVYVYNDNTELPAMGYEADDEGTSQVFSISGDDGYKL--WNMINDPKTEV 568
DB 512 AGIIVNT-----DGLATPMT--STALTTPTFPGLSSVGTQKLVDMVTAHHDSDLGV 562
QY 569 KRNKKEPDKLE-QYPIIDMESFNSKPNVGEDEKIDFKAPDTDELYKEDIIVPAGS 627
DB 563 KIT-----LMLTPQKYTEDEMS-----DF----- 582
QY 628 TSMGPRIIDLKLPVSAAGKNIKSTLVNNGKST 661
DB 583 TSYGPVSNLSFKPDITAPGGINISTON-ANGYTN 615

RESULT 9

AA10941 standard; Protein; 1968 AA.

AC AA10941;

DT 15-APR-1991 (first entry)

DE Mutant protease (del1a137-139/ins9).

KW Mutant protease gene; fermentation; foodstuff; flavouring;

XX Lactic acid bacteria.

OS Lactococcus lactis SK11.

PH Key Location/Qualifiers

FT Peptide 1..187

FT Protein 188..1968

PN EP411715-A.

PD 06-FEB-1991.

PF 02-AUG-1990; 90EP-0202113.

PR 04-AUG-1989; 89NL-0002010.

PA (NEZU-) NED INST ZUIVELONDE.

PI Vos PAJ, Slezien RJ, De Vos WM, Kok J, Venema G, Haandrikman AJ;

DR WP1: 1991-038622/06.

XX N-PSDB; AAQ10871.

PT Mutant protease gene(s) and protease(s) - derived from type I and

PT III protease genes from lactococcal strains, used in fermentation

XX foodstuffs and flavourings

PS Diaclosure; Fig 1(1-7)+5(b); 29pp; English.

XX The wild-type L. lactis SK11 protease gene sequence was determined

CC by the applicant (BP-307011).

CC The mutant protease having new cleavage specificities is

CC obtained by deleting three amino acids (nine bps) and inserting

CC 9 other residues.

CC This mutant may then be used to prepare hybrid proteases,

CC the fusion being between a type I and a type III protease of

CC L. lactis M92 and SK11.

CC The product has modified properties, e.g. thermostability,

CC alkaline/acid pH stability, oxidative stability, autolysis etc.,

CC compared to the parent protease(s). The proteases can be used for

CC preparing products (butter cheese, human and animal foodstuffs)

CC prepared with the aid of lactic acid bacteria.

CC See also AAQ10411-17 and AAQ10870-71.

XX Sequence 1968 AA;

Query Match 20.2%; Score 889.5; DB 12; Length 1968;

Best Local Similarity 25.7%; Pred. No. 0.00082;
Matches 180; Conservative 132; Mismatches 247; Indels 141; Gaps 31;

QY 1 YPV--VLADTSSSEDALNISDEKVAENKEHENIHSAMETSODPFKEKTAIVIKEKVS 58
DB LPVGEIOAKAISOQTGSSILANTVTAA-----TAKQAATDTTAAT----- 65
QY KNPIVDNNTSNEAKIKEENSKNSQGDYTDSPFNKNTENPKEDKVVYIAFPKESGEK 118
DB 66 -NQAIAQLAAKIDYNKLNKVQODIYDVIVQM-SAAPASENGILRTDYSSTAETQOE 123
QY 119 AIKELSLAKTKVLY-----TYDRIFNQSAIETTPDNLDKIQIEGISSVERAOK 168

```

Db      124  TNKVIAQASVKAAVEQVTQGTAGESYGVVNGSTKVRVVDIPKLOIAQVKTVTLAKV 183
Qy      169  VQPMNHNARKEIGVEEADIDYLSINAPFGKNFDRGWNISNIDGTDRYRHKAMRIDDDAK 228
      184  YPPTDAKANSMANVQAQWASNTK-----YKEGTVASVIDSGIDPTHDKMLSDDXD 234
Qy      229  ASMRFKKEDL-KGTDK---NYWLSDKIPHAENVY-VNGSKITVEKYDDGRDYFDPHGMAIA 283
      235  V--KLTKSDVEKFTDVTGHRGYFNSKVYGFYVADNNDTIT---DDKVD--EQGHMIVA 286
Qy      284  GILAGNDEQDIKNPFGIDGIAAPNAQIFSYKMYSD-----AG-SGFAGDETMFHAIEDS 336
      287  GIIANGTGDDPA--KSVVGVAPEAQOLAMKVFNSDSITSGSAGTGTGGSATVVSALIEDS 344
Qy      337  IKNHVDVVS--FTGTGLVGEKYOQAIRALKRAGIPMVVATGNTVATSASSSMDLVAN 395
      345  AKIGADVAMSLGNSNGQTLDEPRLAIVQANANESGTAAVISAQNSGTSSGATBE--VNK 402
Qy      396  NHLKMTDTGNVTRTAHEDAIIVASAKQVVEFDKVI-----GGESFRY--RNIGA 445
      403  DYVGLQDNEMVGSFGTSRGATTVASAEKNTDVTQAVTITDGTGLQIGPETIQLSGHPTG 462
Qy      446  FPDKSKITTNDDGTAPSKLKFVYIGKQODDILGLDKRTIAVMDRI-YTKDLKNAKK 504
      463  SFDQKKFYIVKASGN-----LSKALADYTA-DAKGKIAIVKGEFSPDDKQKQAQ 513
Qy      505  AMDKARAIMVNTVNYNRDNMTLPAWGAEADGTSQVFSISGDDGVKL--WNMINP 562
      514  AA--GAALIIIVT-----DGTATPMT--SIALTTTPFTGLSSVTGQLVDMVTAHP 562
Qy      563  DKTEVKNKKEDFKDLE-QYYPIDMESFNSNKNPNVGEDEKIDFKAPDIDKELYKEDI 621
      563  DDSLGVKIT-----LAMLPNQKYTEDKMS-----DF----- 588
Qy      622  IVPAGSTWGPRIIDLKPDVASRGNIKSTLVANVNGKST 661
      589  -----TSYGPVSNLSFSPDITAPGGINWSTON--NNGYTN 621

```

RESULT 10

AAR10558

ID AAR10558 standard; Protein; 1962 AA.

AAR10558;

15-APR-1991 (first entry)

DB Mutant protease (A137G/K138P/T139P).

XX Mutant protease gene; fermentation; foodstuff; flavouring;

KW lactic acid bacteria.

XX Lactococcus lactis SK11.

XX Key Location/Qualifiers

FT Peptide 1..187

FT Protein /label= signal_peptide

FT Protein 188..1962

FT Protein /label= mature_protein

XX BP111715-A.

XX 06-FEB-1991.

XX 02-AUG-1990; 90BP-0202113.

XX 04-AUG-1989; 89NL-0002010.

XX (NEZU-) NED INST ZUIVELONDE.

XX Vos PAJ, Siezen RJ, De Vos WM, Kok J, Venema G, Haandrikman AJ;

XX WPI, 1991-038622/06.

```

DR      N-PSDB; AAQ10412.
XX      Mutant protease gene(s) and protease(s) - derived from type I and
PT      III protease genes from lactococcal strains, used in fermentation
PT      foodstuffs and flavourings
XX
XX      Disclousure; Fig 1(1-7)+5(b); 29 pp; English.
PS      The mutant protease having new cleavage specificities is
SS      obtained by replacing three amino acids.
CC      This mutant may then be used to prepare hybrid proteases,
CC      the fusion being between a type I and a type III protease of
CC      L.lactis Wg2 and SK11.
CC      The product has modified properties, e.g. thermostability,
CC      alkaline/acid pH stability, oxidative stability, autolysis etc.,
CC      compared to the parent protease(s). The proteases can be used for
CC      preparing products (butter cheese, human and animal foodstuffs)
CC      prepared with the aid of lactic acid bacteria.
CC      See also AAQ10411-17 and AAQ10870-71.
CC
SQ      Sequence 1962 AA;

```

Query Match 20.2%; Score 888.5; DB 12; Length 1962;
 Best Local Similarity 25.8%; Pred. No. 0.00084;
 Matches 179; Conservative 131; Mismatches 249; Indels 135; Gaps 30;

```

Qy      1  YPV-VLADTSSSEDALNISDEKVAENKEHENIHAMETSQDFKEKTAIVIKEVVS 58
      24  LPVEIQAKAIAISQTLGSSLANVTAA-----TAKQAATTTTAA----- 65
Db
Qy      59  KNPIVDNNTSNEAKIKENSNSQGDYTSFVNKNTENPKEDKVVYIAEFKKEGSEK 118
      66  -NQAIAVQLAAKGIDYNNKLVQODIYVDYIVQM--SAAPASENGIIRTDYSSTAIEIOE 123
Db
Qy      119  AIKELSLAKNTKLVY-----TYDRIENGSAIETTPNMLKIOEGISSVERAK 168
      124  TNKVIAQASVKAAVEQVTQGTAGESYGVVNGSTKVRVVDIPKLOIAQVKTVTLAKV 183
Db
Qy      169  VQPMNHNARKEIGVEEADIDYLSINAPFGKNFDRGWNISNIDGTDRYRHKAMRIDDDAK 228
      184  YPPTDAKANSMANVQAQWASNTK-----YKEGTVASVIDSGIDPTHDKMLSDDXD 234
Db
Qy      229  ASMRFKKEDL-KGTDK---NYWLSDKIPHAENVY-VNGSKITVEKYDDGRDYFDPHGMAIA 283
      235  V--KLTKSDVEKFTDVTGHRGYFNSKVYGFYVADNNDTIT---DDKVD--EQGHMIVA 286
Db
Qy      284  GILAGNDEQDIKNPFGIDGIAAPNAQIFSYKMYSDAG--SGFAGDETMFHAIEDSIKNVD 342
      287  GIIANGTGDDPA--KSVVGVAPEAQOLAMKVFNSDSITSGSPGSATVVSALIEDSAKIGAD 344
Db
Qy      345  VVSVSSG--FTGTGLVGEKYOQAIRALKRAGIPMVVATGNTVATSASSSMDLVANHLKMT 401
      345  VLNMSLGSNSNGQTLDEPRLAIVQANANESGTAAVISAQNSGTSSGATBE--VNDYVQLQ 402
Db
Qy      402  DTGNVTRTAHEDAIIVASAKQVVEFDKVI-----GGESFRY--RNIGAEPDKSK 451
      403  DNEWVSGPSTRGATTVASAEKNTDVTQAVTITDGTGLQIGPETIQLSGHPTGSPDQKK 462
Db
Qy      452  ITTNEDEGKAPSKLKFVYIGKQODDILGLDKRTIAVMDRI-YTKDLKNAFKKAMDGA 510
      463  FYIVKASGN-----LSKALADYTA-DAKGKIAIVKGEFSPDDKQKQAQDA--GA 511
Db
Qy      511  RAIMVNTVNYNRDNMTLPAWGAEADGTSQVFSISGDDGVKL--WNMINPDKTEV 568
      512  AGLIIVT-----DGTATPMT--SIALTTTPFTGLSSVTGQLVDMVTAHDDDSLGV 562
Qy      569  KRNKKEDFKDLE-QYYPIDMESFNSNKNPNVGEDEKIDFKAPDIDKELYKEDIIVPAGS 627
      563  KIT-----LAMLPNQKYTEDKMS-----DF----- 582
Db
Qy      628  TSGGPRIIDLKPDVASRGNIKSTLVANVNGKST 661
      583  TSYGPVSNLSFSPDITAPGGINWSTON--NNGYTN 615
Db

```



```

XX WP1, 1991-036622/06.
DR N-PSDB; AA010414.
XX Mutant protease gene(s) and protease(s) - derived from type I and
PT III protease genes from lactococcal strains, used in fermentation
XX foodstuffs and flavourings
XX Disclosure: Fig 1(1-7)+(a); 29 pp; English.
XX
XX The mutant protease K138n having new cleavage specificities is
CC obtained by carrying out single amino acid substitutions.
CC This mutant may then be used to prepare hybrid proteases,
CC the fusion being between a type I and a type III protease of
CC L.lactis Wg2 and Skil1.
CC The product has modified properties, e.g. thermostability,
CC alkaline/acid pH stability, oxidative stability, autolysis etc.,
CC compared to the parent protease(s). The proteases can be used for
CC preparing products (butter cheese, human and animal foodstuffs)
CC prepared with the aid of lactic acid bacteria.
CC See also AA010411-17 and AA010870-71.
XX
XX Sequence 1962 AA;
SQ
Query Match 20.1%; Score 885.5; DB 12; Length 1962;
Best Local Similarity 25.6%; Pred. No. 0.00094;
Matches 178; Conservative 132; Mismatches 249; Indels 135; Gaps 30.
OY 1 YPV--VLADTSSSDALNIDSEKYAENKEKEHNHSAETSODPEKEKTAIVIKEKVV58
DB 24 LPVSGIOAKKAIISOOTLGSSILANTYAA-----TAKQATDTTAAT-----65
OY 59 KNPVIDNNTSNEEAKIKEKSNKSGDYTDSFVNKNTENPKEDKVVYIAEFKDESEK118
DB 66 -NQAIAATQAAAGIDYNKLNKRVQOQDIYDVIVQM-SAAPASENGILRTDYESTAIQOE123
OY 119 AIKESSLKNTVLY-----TYDRIFNGSALETTPDNLDIKIQLEGISVERAK168
DB 124 TNKVIAAASVAAAEQVYTOQTAGESYGVVNGFSTKRVVDIPLKIQIAGKVTYLLKV183
OY 169 VOPMNAHAKKEIGVEEALDYLSINAPFCGNDFGSGMVISNIDTGDVYHKKMRIDDDAK228
DB 184 YFPTAKANSMAVQAVASNYK-----YKGEYIVSVISDGIIDPHKKMRISDDKD234
OY 229 ASMRPEKEDL-KGTDK--NYWLSDKIPHAIFY-YNGKITYEKYDGDGRDYEPHGMHIA283
DB 235 V--KLTSGDVEKFTDTVHKGRFNSKVPYGFNYADNNDIT----DKYD--EQGMVYA286
OY 284 GILAGNTEBODIKAFNNGIDGILAPNAQISYKMYSDAG-SGPAGDETMFAHIDSTIKHVD342
DB 287 GIIAGNGTGDPPA--KSVVGVAPEAQILAMKVFNSDTSANTGSAITVSAIEDSAKIGAD344
OY 343 VVSVSSG-FTGTGLGVEKYMOWAIRLRKGIIMVVAATGVATYSASSSSMDVLAAHHLMKT401
DB 345 VLANSLSGNSGQTLDEPFLAAVQANBSGTAIVISAGNSGTSGBATBS--VNKDYYGQ402
OY 402 DTGNVTRTPAAHEDAIAYASAKQOTVEFDKVIN-----GGESEFY--RNIGAFEDSK451
DB 403 DNEWVGSGTSGATTTVASAENTVITQAVTTITDGTGLQGEFTITQLSHDPTSGFDQK462
OY 452 ITTNEDGTAPSKLKFVYIGKQODDLIGLDIRGKIAYMDRI-YTKDLKNAEFKAMDKGA510
DB 453 FYIVXDAAGN-----LSKGLALDYTA-DAKGIKAIIVKREGEFDFDKQKVAQAA--GA511
OY 511 RAIMVNVVYVYNRNDMTPLPMAGVEADBGTSQVFSISGDDGVTL--WNMNPDKTEY568
DB 512 AGLIIVNT-----DGTATPMT--SIALTTTFPFGLSSVTGQKLVDMVTAHPDLSIGV562
OY 569 KRNKEDFDKLE-QYYPIDMESFNSKNKPNVDEKEIDPKFAPDTDKELYKEDIIVPAGS627
DB 563 KIT-----LAMLPNOKYTEDDKMS-----DP-----582
OY 628 TSWGPRIDLLKPKDVSAPGKNIKSTLVNINGSKST 661

```

Query	Subject	Score	Expect	Ident	Length	Gap
1	AA10561 standard; Protein, 1962 AA.	20.1%	Score 885.5; DB 12; Length 1962; Bact local similarity 25.6%; Pred. No. 0.00094; Matches 178; Conservative 132; Mismatches 249; Indels 135; Gaps 30.	178	1962	30
2	AA10561; 15-APR-1991 (first entry)	20.1%	Score 885.5; DB 12; Length 1962; Bact local similarity 25.6%; Pred. No. 0.00094; Matches 178; Conservative 132; Mismatches 249; Indels 135; Gaps 30.	178	1962	30
3	AA10561; Mutant protease (N166D).	20.1%	Score 885.5; DB 12; Length 1962; Bact local similarity 25.6%; Pred. No. 0.00094; Matches 178; Conservative 132; Mismatches 249; Indels 135; Gaps 30.	178	1962	30
4	AA10561; Mutant protease gene, fermentation; foodstuff, flavouring; lactic acid bacteria.	20.1%	Score 885.5; DB 12; Length 1962; Bact local similarity 25.6%; Pred. No. 0.00094; Matches 178; Conservative 132; Mismatches 249; Indels 135; Gaps 30.	178	1962	30
5	AA10561; Lactococcus lactis SK11.	20.1%	Score 885.5; DB 12; Length 1962; Bact local similarity 25.6%; Pred. No. 0.00094; Matches 178; Conservative 132; Mismatches 249; Indels 135; Gaps 30.	178	1962	30
6	AA10561; Key	20.1%	Score 885.5; DB 12; Length 1962; Bact local similarity 25.6%; Pred. No. 0.00094; Matches 178; Conservative 132; Mismatches 249; Indels 135; Gaps 30.	178	1962	30
7	AA10561; Peptide	20.1%	Score 885.5; DB 12; Length 1962; Bact local similarity 25.6%; Pred. No. 0.00094; Matches 178; Conservative 132; Mismatches 249; Indels 135; Gaps 30.	178	1962	30
8	AA10561; Protein	20.1%	Score 885.5; DB 12; Length 1962; Bact local similarity 25.6%; Pred. No. 0.00094; Matches 178; Conservative 132; Mismatches 249; Indels 135; Gaps 30.	178	1962	30
9	AA10561; EP411715-A.	20.1%	Score 885.5; DB 12; Length 1962; Bact local similarity 25.6%; Pred. No. 0.00094; Matches 178; Conservative 132; Mismatches 249; Indels 135; Gaps 30.	178	1962	30
10	AA10561; 06-FEB-1991.	20.1%	Score 885.5; DB 12; Length 1962; Bact local similarity 25.6%; Pred. No. 0.00094; Matches 178; Conservative 132; Mismatches 249; Indels 135; Gaps 30.	178	1962	30
11	AA10561; 02-AUG-1990; 90EP-0202113.	20.1%	Score 885.5; DB 12; Length 1962; Bact local similarity 25.6%; Pred. No. 0.00094; Matches 178; Conservative 132; Mismatches 249; Indels 135; Gaps 30.	178	1962	30
12	AA10561; 04-AUG-1989; 89NL-0002010.	20.1%	Score 885.5; DB 12; Length 1962; Bact local similarity 25.6%; Pred. No. 0.00094; Matches 178; Conservative 132; Mismatches 249; Indels 135; Gaps 30.	178	1962	30
13	AA10561; PR	20.1%	Score 885.5; DB 12; Length 1962; Bact local similarity 25.6%; Pred. No. 0.00094; Matches 178; Conservative 132; Mismatches 249; Indels 135; Gaps 30.	178	1962	30
14	AA10561; (NEZU-) NED INST ZUIVELONDE.	20.1%	Score 885.5; DB 12; Length 1962; Bact local similarity 25.6%; Pred. No. 0.00094; Matches 178; Conservative 132; Mismatches 249; Indels 135; Gaps 30.	178	1962	30
15	AA10561; Vos PAJ, Siezen RJ, De Vos WM, Kok J, Venema G, Haandrikman AJ;	20.1%	Score 885.5; DB 12; Length 1962; Bact local similarity 25.6%; Pred. No. 0.00094; Matches 178; Conservative 132; Mismatches 249; Indels 135; Gaps 30.	178	1962	30
16	AA10561; WPI; 1991-038622/06.	20.1%	Score 885.5; DB 12; Length 1962; Bact local similarity 25.6%; Pred. No. 0.00094; Matches 178; Conservative 132; Mismatches 249; Indels 135; Gaps 30.	178	1962	30
17	AA10561; DR N-PSDB; AAQ10415.	20.1%	Score 885.5; DB 12; Length 1962; Bact local similarity 25.6%; Pred. No. 0.00094; Matches 178; Conservative 132; Mismatches 249; Indels 135; Gaps 30.	178	1962	30
18	AA10561; Mutant protease gene(s) and protease(s) - derived from type I and III protease genes from lactococcal strains, used in fermentation foodstuffs and flavourings	20.1%	Score 885.5; DB 12; Length 1962; Bact local similarity 25.6%; Pred. No. 0.00094; Matches 178; Conservative 132; Mismatches 249; Indels 135; Gaps 30.	178	1962	30
19	AA10561; PT	20.1%	Score 885.5; DB 12; Length 1962; Bact local similarity 25.6%; Pred. No. 0.00094; Matches 178; Conservative 132; Mismatches 249; Indels 135; Gaps 30.	178	1962	30
20	AA10561; Disclousure; Fig 1(1-7)+5(a); 29 pp; English.	20.1%	Score 885.5; DB 12; Length 1962; Bact local similarity 25.6%; Pred. No. 0.00094; Matches 178; Conservative 132; Mismatches 249; Indels 135; Gaps 30.	178	1962	30
21	AA10561; PS	20.1%	Score 885.5; DB 12; Length 1962; Bact local similarity 25.6%; Pred. No. 0.00094; Matches 178; Conservative 132; Mismatches 249; Indels 135; Gaps 30.	178	1962	30
22	AA10561; CC	20.1%	Score 885.5; DB 12; Length 1962; Bact local similarity 25.6%; Pred. No. 0.00094; Matches 178; Conservative 132; Mismatches 249; Indels 135; Gaps 30.	178	1962	30
23	AA10561; CC	20.1%	Score 885.5; DB 12; Length 1962; Bact local similarity 25.6%; Pred. No. 0.00094; Matches 178; Conservative 132; Mismatches 249; Indels 135; Gaps 30.	178	1962	30
24	AA10561; CC	20.1%	Score 885.5; DB 12; Length 1962; Bact local similarity 25.6%; Pred. No. 0.00094; Matches 178; Conservative 132; Mismatches 249; Indels 135; Gaps 30.	178	1962	30
25	AA10561; CC	20.1%	Score 885.5; DB 12; Length 1962; Bact local similarity 25.6%; Pred. No. 0.00094; Matches 178; Conservative 132; Mismatches 249; Indels 135; Gaps 30.	178	1962	30
26	AA10561; CC	20.1%	Score 885.5; DB 12; Length 1962; Bact local similarity 25.6%; Pred. No. 0.00094; Matches 178; Conservative 132; Mismatches 249; Indels 135; Gaps 30.	178	1962	30
27	AA10561; CC	20.1%	Score 885.5; DB 12; Length 1962; Bact local similarity 25.6%; Pred. No. 0.00094; Matches 178; Conservative 132; Mismatches 249; Indels 135; Gaps 30.	178	1962	30
28	AA10561; CC	20.1%	Score 885.5; DB 12; Length 1			

```

Db      124 TNKVIAAQASVKAALVEQVTOQTAGESTGYVNGFSTKRVVDIPKQLQIAGVKTVTLAKV 183
Qy      169 VQPMNHARKKEIGVEBEAIDYLKSNAPFGKNPDRGKGVNISIDTGYTRHAKMRIDDDAK 228
Db      184 YYPIDAKANSMAVNAQAVMSNYK-----YKCEGVVSVYIDGIDPTKMDMLSDDKD 234
Qy      229 ASMRFKKEDL-KGTDK--NYWLSDKIPHAENV-YNGSKITVEKYDGRDYPDPHGMHIA 283
Db      235 V--KLTKSDVKEKFTDVYKGRYFNSKVPYGFNVADNDITL---DDKVD--EQHGHHVA 286
Qy      284 GILAGNTEQDIKNFNGIDGIAFPAQIFSYKMYSDAG--SGFAGDETFPHAIEDSIKXNV 342
Db      287 GILGANGTGDDPA--KSVVGVAPEAQLAMKVFNSNSTSANTGSATVVSALIEDSAKIGAD 344
Qy      343 VSVSSGF-TGTGLVGEKYMQAIRALRKAGIPMVVATGNTVTSASSSMDLVANNHLKMT 401
Db      345 VLNMSLGSDSGNQTLDEBELAAVONANESGTAIVISAGNSGTSGSATEG--VKKDYVGLQ 402
Qy      402 DTGNVTRTAHEDAIIVASAKNQTFEEDKVN-----GGSFPKY--RNIGAFDPKSK 451
Db      403 DNEWVSGPSTSRGATVYASAEVNTDVIITQAVTITDGTGLQIGPETIQLSHDFTGSPQOK 462
Qy      452 ITTNEDGTAPSKAKFYVYIGGDODLIGDLRGKIAVMRI-YTKDLKNAFKKAMDKGA 510
Db      463 FYIVKQASGN-----LSKGLADLYTA--DAKGIKAIIVKRGESFPDKOKYQAQA--GA 511
Qy      511 RAIMVNTVTVYVNRDNTLPAWGYBADEGTSQVFSISGDDGVKL--MMNINPDKKTREY 568
Db      512 AGLIIVNT-----DGTATPMT--SIALTTTPTFGLSVTTGQKLVDTWTAHPDSDSLGV 562
Qy      569 KRNKKEPDFOKLE-QYYPIDMESFNSKNPNVGEKEIDPFKAPDTELKELYKEDIIVPAGS 627
Db      563 KIT-----LAMLPNQKYTEDKMS-----DF-----582
Qy      628 TSWGPRIDLLKPDVSAFGKNIKSTLVINGKST 661
Db      583 TSYGPAVNSLSFKPDITAPGGINWSTON--NNGXTN 615

RESULT 14
AARI0563
ID AARI0563 standard; Protein; 1962 AA.
XX
AC AARI0563;
XX
DT 15-Apr-1991 (first entry)
XX
DE Mutant protease (K748T).
XX
KM Mutant protease gene; fermentation; foodstuff; flavouring;
XX
KM lactic acid bacteria.
XX
OS Lactococcus lactis SK11.
XX
FH Key Location/Qualifiers
FT Peptide 1..187
FT Protein /label= signal_peptide
FT Protein 188..1962
FT Protein /label= mature_protein
XX
PN BP411715-A.
PD 06-FEB-1991.
XX
PF 02-AUG-1990; 90EP-0202113.
XX
PR 04-AUG-1989; 89NL-0002010.
XX
PA (NEZU-) NED INST ZUIVELONDE.
XX
PI Vos PAJ, Slezzen RJ, De Vos WM, Kok J, Venema G, Haandrikman AJ;

```

```

DR WPI; 1991-038622/06.
DR N-PSDB; AAQ10417.
XX
PT Mutant protease gene(s) - derived from type I and
PT III protease genes from lactococcal strains, used in fermentation
PT foodstuffs and flavourings
XX
PS Diaclosure; Fig 1(1-7)+5(a); 29 pp; English.
XX
CC The mutant protease K748T having new cleavage specificities is
CC obtained by carrying out single amino acid substitutions.
CC This mutant may then be used to prepare hybrid proteases,
CC the fusion being between a type I and a type III protease of
CC L. lactis Wg2 and SK11.
CC The product has modified properties, e.g. thermostability,
CC alkaline/acid pH stability, oxidative stability, autolysis etc.,
CC compared to the parent protease(s). The proteases can be used for
CC preparing products (butter cheese, human and animal foodstuffs)
CC prepared with the aid of lactic acid bacteria.
CC See also AAQ10411-17 and AAQ10870-71.
XX
SQ Sequence 1962 AA.

Query Match 20.1%; Score 885.5; DB 12; Length 1962;
Best Local Similarity 25.6%; Pred. No. 0.00094;
Matches 178; Conservative 132; Mismatches 249; Indels 135; Gaps 30;

Qy 1 YPV--VLADTSSEEDALNIDSKERVAENKEKHENHSAMETSODPFKEKTAIVKEKSVS 58
Db 24 LPVGEIQAQKAIQOQTGLSSLVNTVTA-----YAKAAADTDTAATV-----65
Qy 59 KNPVIDNNTSNEBAKIEKSNKSGQDYTSPFNKATENPKCKDKVYVIAFPKDESGEK 118
Db 66 -NQAIAQLAKAIDVYKLNKVKQOQDIYDVYIQW-SAAPASNGILRTYSSATRIQOE 123
Qy 119 AIEELSLKNTKVLV-----TYDRIENGSALETTDPNDLKIKQIEGISSVERAQK 168
Db 124 TNKVIAAQASVKAALVEQVTOQTAGESTGYVNGFSTKRVVDIPKQLQIAGVKTVTLAKV 183
Qy 169 VQPMNHARKKEIGVEBEAIDYLKSNAPFGKNPDRGKGVNISIDTGYTRHAKMRIDDDAK 228
Db 184 YYPIDAKANSMAVNAQAVMSNYK-----YKCEGVVSVYIDGIDPTKMDMLSDDKD 234
Qy 229 ASMRFKKEDL-KGTDK--NYWLSDKIPHAENV-YNGSKITVEKYDGRDYPDPHGMHIA 283
Db 235 V--KLTKSDVKEKFTDVYKGRYFNSKVPYGFNVADNDITL---DDKVD--EQHGHHVA 286
Qy 284 GILAGNTEQDIKNFNGIDGIAFPAQIFSYKMYSDAG--SGFAGDETFPHAIEDSIKXNV 342
Db 287 GILGANGTGDDPA--KSVVGVAPEAQLAMKVFNSNSTSANTGSATVVSALIEDSAKIGAD 344
Qy 343 VSVSSGF-FGTGTLVGEKYMQAIRALRKAGIPMVVATGNTVTSASSSMDLVANNHLKMT 401
Db 345 VLNMSLGSDSGNQTLDEBELAAVONANESGTAIVISAGNSGTSGSATEG--VKKDYVGLQ 402
Qy 402 DTGNVTRTAHEDAIIVASAKNQTFEEDKVN-----GGSFPKY--RNIGAFDPKSK 451
Db 403 DNEWVSGPSTSRGATVYASAEVNTDVIITQAVTITDGTGLQIGPETIQLSHDFTGSPQOK 462
Qy 452 ITTNEDGTAPSKAKFYVYIGGDODLIGDLRGKIAVMRI-YTKDLKNAFKKAMDKGA 510
Db 463 FYIVKQASGN-----LSKGLADLYTA--DAKGIKAIIVKRGESFPDKOKYQAQA--GA 511
Qy 511 RAIMVNTVTVYVNRDNTLPAWGYBADEGTSQVFSISGDDGVKL--MMNINPDKKTREY 568
Db 512 AGLIIVNT-----DGTATPMT--SIALTTTPTFGLSVTTGQKLVDTWTAHPDSDSLGV 562
Qy 569 KRNKKEPDFOKLE-QYYPIDMESFNSKNPNVGEKEIDPFKAPDTELKELYKEDIIVPAGS 627
Db 563 KIT-----LAMLPNQKYTEDKMS-----DF-----582
Qy 628 TSWGPRIDLLKPDVSAFGKNIKSTLVINGKST 661

```

